Московский государственный университет им. М. В. Ломоносова
Физический факультет
Кафедра компьютерных методов физики

М. Л. Сердобольская

Методы функционального анализа в задачах редукции

2014 г.

СПИСОК ОБОЗНАЧЕНИЙ

$\mathbb{R}$ — множество (поле) действительных чисел, $\mathbb{R}_+$ — множество неотрицательных действительных чисел, $\mathbb{N}$ — множество натуральных чисел;

$\mathcal{R}$, $\widetilde{\mathcal{R}}$ и т. п. — действительные сепарабельные гильбертовы пространства бесконечной (если иное не оговорено особо) размерности;

$\{x_n\}_{n=\overline{1,\infty}}$ — последовательность элементов $x_1, \ldots, x_n, \ldots$;

$\overline{\mathcal{M}} = \bigl\{x \in \mathcal{R}\colon\ x = \lim\limits_{n\to\infty} x_n,\ \text{где } \{x_n\}_{n=\overline{1,\infty}} \subset \mathcal{M}\bigr\}$ — замыкание множества $\mathcal{M}$;

$\mathcal{M}^{\perp} = \{x \in \mathcal{R}\colon\ (x, z) = 0 \text{ для любого } z \in \mathcal{M}\}$ — ортогональное дополнение множества $\mathcal{M}$;

$\mathbb{L}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ — множество линейных операторов, действующих из пространства $\mathcal{R}$ в пространство $\widetilde{\mathcal{R}}$;

$\mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ — множество замкнутых плотно определённых операторов, действующих из пространства $\mathcal{R}$ в пространство $\widetilde{\mathcal{R}}$

$\mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ — банахово пространство ограниченных всюду определённых линейных операторов, действующих из пространства $\mathcal{R}$ в пространство $\widetilde{\mathcal{R}}$;

$\mathbb{CO}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ — банахово пространство ограниченных всюду определённых компактных операторов, действующих из пространства $\mathcal{R}$ в пространство $\widetilde{\mathcal{R}}$ (линейное подпространство пространства $\mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$);

$\mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ — гильбертово пространство всюду определённых линейных операторов Гильберта–Шмидта, действующих из пространства $\mathcal{R}$ в пространство $\widetilde{\mathcal{R}}$;

$\mathcal{D}(A) = \{x \in \mathcal{R}\colon\ \text{существует } Ax\}$ — область определения линейного оператора A, действующего из пространства $\mathcal{R}$ в пространство $\widetilde{\mathcal{R}}$;

$\mathcal{N}(A) = \{x \in \mathcal{D}(A)\colon\ Ax = 0\}$ — нуль-пространство (ядро) линейного оператора A, действующего из пространства $\mathcal{R}$ в пространство $\widetilde{\mathcal{R}}$;

$\mathcal{R}(A) = \{y \in \widetilde{\mathcal{R}}\colon\ \text{существует } x \in \mathcal{D}(A), \text{ для которого } y = Ax\}$ — пространство значений (образ) линейного оператора A, действующего из пространства $\mathcal{R}$ в пространство $\widetilde{\mathcal{R}}$;

$\mathrm{E}\xi$ — математическое ожидание случайного элемента (случайной величины) ξ.

В предельных переходах типа $x_n \to x$ или $\lim x_n = x$ мы не всегда будем писать условие $n \to \infty$ в тех случаях, когда это не может вызвать недоразумений. То же замечание относится и к последовательностям $\{x_n\}$ элементов с одним индексом.

ГЛАВА 1
СЛУЧАЙНЫЕ ЭЛЕМЕНТЫ В ГИЛЬБЕРТОВЫХ ПРОСТРАНСТВАХ

1.1. Понятие случайного элемента. Пусть $(\Omega, \mathcal{F}, P)$ — некоторое вероятностное пространство. Основные понятия и некоторые полезные теоремы теории вероятностей можно найти в приложении А.

Определение. Элементом ξ, *случайным в слабом смысле*, принимающим значения в гильбертовом пространстве $\mathcal{R}$ называется правило $\xi\colon \Omega \mapsto \mathcal{R}$, однозначным образом сопоставляющее элементарному исходу $\omega \in \Omega$ элемент $\xi(\omega) \in \mathcal{R}$ так, что для любого фиксированного $x \in \mathcal{R}$ скалярное произведение $(\xi(\omega), x)$, рассматриваемое как функция из Ω в $\mathbb{R}$, есть случайная величина.

В дальнейшем будем называть такой элемент «случайным элементом в $\mathcal{R}$», опуская слова «в слабом смысле». Непосредственно из определения вытекает, что линейная комбинация $\lambda_1\xi_1 + \lambda_2\xi_2$ случайных элементов (действующих из одного и того же Ω в одно и то же $\mathcal{R}$) с постоянными коэффициентами $\lambda_{1,2} \in \mathbb{R}$ есть случайный элемент. Кроме того, $\|\xi\|^2$ также есть случайный элемент. Это следует из того, что[1)]

$$(\lambda_1\xi_1 + \lambda_2\xi_2, x) = \lambda_1(\xi_1, x) + \lambda_2(\xi_2, x) \quad \text{и} \quad \|\xi\|^2 = \sum_{i=1}^{\infty}(\xi, e_i)^2 = \lim_{n\to\infty}\sum_{i=1}^{n}(\xi, e_i)^2$$

суть случайные величины как линейная комбинация и предел последовательности случайных величин соответственно (см. теорему А.1 приложения А). В последних равенствах x — произвольный элемент из $\mathcal{R}$ и $e_1, e_2, \ldots$ — произвольный ортонормированный базис (ОНБ) пространства $\mathcal{R}$; разумеется, численное значение любой реализации $\|\xi\|^2(\omega)$, $\omega \in \Omega$, не зависит от базиса.

Пусть R — линейный оператор из $\mathcal{R}$ в $\widetilde{\mathcal{R}}$. Если $R \in \mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, то соответствие $R\xi\colon \Omega \mapsto \widetilde{\mathcal{R}}$, заданное формулой $(R\xi)(\omega) = R \cdot \xi(\omega)$, есть случайная величина. В самом деле, для любого $y \in \widetilde{\mathcal{R}}$ скалярное произведение $(R\xi, y) = (\xi, R^*y)$ есть случайная величина. Написав последнее равенство, мы воспользовались тем, что

[1)] Всюду далее, когда мы пишем равенства и неравенства для случайных величин, мы подразумеваем, что они имеют место для любого $\omega \in \Omega$, точнее — с вероятностью единица.

для оператора $R \in \mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ существует всюду определённый сопряжённый оператор R^*.

1.2. Моменты случайных элементов. Пусть случайная величина $\|\xi\|^2$ имеет математическое ожидание: существует (вообще говоря, несобственный) интеграл Лебега–Стилтьеса

$$\mathrm{E}\|\xi\|^2 = \int_{\mathbb{R}} t\, dF(t) < \infty,$$

где $F(\cdot)$ — функция распределения случайной величины $\|\xi\|^2$. Напишем цепочку равенств

$$\mathrm{E}\|\xi\|^2 = \mathrm{E}\sum_{i=1}^{\infty}(\xi, e_i)^2 \overset{?}{=} \sum_{i=1}^{\infty}\mathrm{E}(\xi, e_i)^2$$

и выясним, справедливо ли равенство, помеченное вопросительным знаком. В самом деле, возможность поменять местами знак математического ожидания и знак суммы ряда — это возможность применить равенство $\mathrm{E}(\lim \alpha_n) = \lim \mathrm{E}\alpha_n$. В данном случае такая замена предельных переходов правомерна. Это вытекает из теоремы A.2 о сходимости математических ожиданий. Мы можем написать

$$\|\xi\|^2 = \lim_{n\to\infty}\sum_{i=1}^{n}(\xi, e_i)^2, \qquad \sum_{i=1}^{n}(\xi, e_i)^2 \leqslant \|\xi\|^2,$$

следовательно,

$$\mathrm{E}\|\xi\|^2 = \lim_{n\to\infty}\mathrm{E}\sum_{i=1}^{n}(\xi, e_i)^2 = \lim_{n\to\infty}\sum_{i=1}^{n}\mathrm{E}(\xi, e_i)^2 = \sum_{i=1}^{\infty}\mathrm{E}(\xi, e_i)^2. \tag{1.1}$$

Определение. Случайный элемент ξ называется *гильбертовым случайным элементом*, если $\mathrm{E}\|\xi\|^2 < \infty$.

Пусть при любом $x \in \mathcal{R}$ существует математическое ожидание случайной величины (ξ, x). Следовательно, всюду на $\mathcal{R}$ определен (очевидно, линейный) функционал $f(x) = \mathrm{E}(\xi, x)$. Если этот функционал ограничен, то по теореме Рисса об общем виде линейного функционала в гильбертовом пространстве найдется единственный (неслучайный) элемент $\mathrm{E}\xi \in \mathcal{R}$ такой, что $f(x) = (\mathrm{E}\xi, x)$ для любого $x \in \mathcal{R}$. В результате мы получаем следующее определение.

Определение. Элемент $\mathrm{E}\xi \in \mathcal{R}$, заданный равенством

$$\mathrm{E}(\xi, x) = (\mathrm{E}\xi, x), \qquad x \in \mathcal{R}, \tag{1.2}$$

называется *математическим ожиданием случайного элемента* ξ.

Заметим, что если ξ — гильбертов случайный элемент, то

$$|\mathrm{E}(\xi,x)| \leqslant \mathrm{E}|(\xi,x)| \leqslant \mathrm{E}\sqrt{\|\xi\|^2\|x\|^2} \leqslant \sqrt{\mathrm{E}(\|\xi\|^2\|x\|^2)} = \|x\|\sqrt{\mathrm{E}\|\xi\|^2}, \qquad (1.3)$$

где мы воспользовались неравенством $|\mathrm{E}\alpha| \leqslant \mathrm{E}|\alpha|$, неравенством Коши–Буняковского, а также неравенством $\mathrm{E}\sqrt{\alpha} \leqslant \sqrt{\mathrm{E}\alpha}$, которое эквивалентно неравенству $\mathrm{D}\sqrt{\alpha} = \mathrm{E}\alpha - \mathrm{E}^2\sqrt{\alpha} \geqslant 0$. Таким образом,

$$\|\mathrm{E}\xi\| = \|f\| = \sup_{\|x\|=1} |\mathrm{E}(\xi,x)| \leqslant \sqrt{\mathrm{E}\|\xi\|^2} < \infty,$$

в результате, если существует $\mathrm{E}\|\xi\|^2$, то существует и $\mathrm{E}\xi \in \mathcal{R}$, причем имеет место неравенство $\|\mathrm{E}\xi\|^2 \leqslant \mathrm{E}\|\xi\|^2$.

Пусть существуют $\mathrm{E}\xi, \mathrm{E}\xi_{1,2}$. Из равенств

$$\mathrm{E}(\lambda_1\xi_1 + \lambda_2\xi_2, x) = \lambda_1\mathrm{E}(\xi_1,x) + \lambda_2\mathrm{E}(\xi_2,x) = (\lambda_1\mathrm{E}\xi_1 + \lambda_2\mathrm{E}\xi_2, x),$$
$$\mathrm{E}(R\xi, y) = \mathrm{E}(\xi, R^*y) = (\mathrm{E}\xi, R^*y) = (R\,\mathrm{E}\xi, y),$$

справедливых для любых $x \in \mathcal{R}$, $y \in \widetilde{\mathcal{R}}$, любых действительных чисел $\lambda_{1,2}$ и любого оператора $R \in \mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, следует, что

$$\mathrm{E}(\lambda_1\xi_1 + \lambda_2\xi_2) = \lambda_1\mathrm{E}\xi_1 + \lambda_2\mathrm{E}\xi_2, \qquad \mathrm{E}R\xi = R\,\mathrm{E}\xi.$$

Пусть ξ — случайный элемент в $\mathcal{R}$ и η — случайный элемент в $\widetilde{\mathcal{R}}$. Пусть существуют математические ожидания $\mathrm{E}\xi = 0$ и $\mathrm{E}\eta = 0$[2)]. Предположим, что для каждого $x \in \mathcal{R}$ и каждого $y \in \widetilde{\mathcal{R}}$ существует $\mathrm{E}(\xi,x)(\eta,y)$, причём найдется константа C такая, что выполнено неравенство

$$|\mathrm{E}(\xi,x)(\eta,y)| \leqslant C\cdot\|x\|\cdot\|y\|, \qquad x \in \mathcal{R}, \quad y \in \widetilde{\mathcal{R}}. \qquad (1.4)$$

Заметим, что если ξ и η — гильбертовы случайные элементы, то в силу цепочки неравенств, аналогичной (1.3), мы имеем

$$|\mathrm{E}(\xi,x)(\eta,y)| \leqslant \sqrt{\|\mathrm{E}\xi\|^2}\sqrt{\|\mathrm{E}\eta\|^2}\cdot\|x\|\cdot\|y\|,$$

таким образом, в этом случае неравенство (1.4) выполнено. При условии (1.4) зафиксируем $x \in \mathcal{R}$, тогда $\mathrm{E}(\xi,x)(\eta,\cdot)$ — ограниченный линейный функционал, определённый на всём пространстве $\widetilde{\mathcal{R}}$. Следовательно, найдется единственный

[2)] Мы не вводим, если в этом нет необходимости, специальных символов для нулевых элементов в разных пространствах, используя единое обозначение 0. Так, $\mathrm{E}\xi = 0$ — нулевой элемент в $\mathcal{R}$, а $\mathrm{E}\eta = 0$ — нулевой элемент в $\widetilde{\mathcal{R}}$.

элемент $u_x \in \widetilde{\mathcal{R}}$ такой, что $E(\xi, x)(\eta, y) = (u_x, y)$ для любого $y \in \widetilde{\mathcal{R}}$. Можно сказать, что соответствие $x \mapsto u_x$ задаёт оператор, который мы обозначим через $\Sigma_{\xi\eta}$:

$$(\Sigma_{\xi\eta} x, y) = \mathrm{E}(\xi, x)(\eta, y), \qquad x \in \mathcal{R}, \quad y \in \widetilde{\mathcal{R}}. \tag{1.5}$$

Аналогично, всюду на $\widetilde{\mathcal{R}}$ можно задать оператор $\Sigma_{\eta\xi}$ со значениями в $\mathcal{R}$ такой, что

$$(x, \Sigma_{\eta\xi} y) = \mathrm{E}(\xi, x)(\eta, y), \qquad x \in \mathcal{R}, \quad y \in \widetilde{\mathcal{R}}. \tag{1.6}$$

Очевидно, что данные формулы задают линейные операторы, причем $\Sigma_{\eta\xi} = \Sigma_{\xi\eta}^*$.

Определение. Операторы $\Sigma_{\xi\eta} \in \mathbb{L}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ и $\Sigma_{\eta\xi} \in \mathbb{L}(\widetilde{\mathcal{R}} \mapsto \mathcal{R})$, заданные равенствами (1.5) и (1.6), называются *взаимными ковариационными операторами случайных элементов* ξ и η.

Определение. Будем говорить, что случайный элемент ξ со значениями в $\mathcal{R}$ и случайный элемент η со значениями в $\widetilde{\mathcal{R}}$ *некоррелированы*, если $\Sigma_{\eta\xi}$ есть нулевой оператор из $\mathcal{R}$ в $\widetilde{\mathcal{R}}$, т. е. $\Sigma_{\xi\eta} x = 0$ (нулевой элемент в $\widetilde{\mathcal{R}}$) для всех $x \in \mathcal{R}$.

В силу равенства $\Sigma_{\eta\xi} = \Sigma_{\xi\eta}^*$ для некоррелированных ξ и η их взаимный ковариационный оператор $\Sigma_{\eta\xi}$ (действующий из $\widetilde{\mathcal{R}}$ в $\mathcal{R}$) также есть нулевой оператор[3)].

Очевидно, что если ξ и η некоррелированы, то $R\xi$ и η также некоррелированы (здесь R — ограниченный всюду определённый линейный оператор):

$$\mathrm{E}(R\xi, x)(\eta, y) = \mathrm{E}(\xi, R^* x)(\eta, y) = (\Sigma_{\xi\eta} R^* x, y) = 0$$

для всех x и y из надлежащих пространств.

Если положить $\eta = \xi$, то мы получим следующее определение.

Определение. Оператор $\Sigma_{\xi\xi} \in \mathbb{L}(\mathcal{R} \mapsto \mathcal{R})$, заданный равенством

$$(x, \Sigma_{\xi\xi} z) = \mathrm{E}(\xi, x)(\xi, z), \qquad x, z \in \mathcal{R}, \tag{1.7}$$

называется *ковариационным оператором случайного элемента* ξ.

Очевидно, что $\Sigma_{\xi\xi} = \Sigma_{\xi\xi}^*$ и $\Sigma_{\xi\xi} \geqslant 0$, поскольку $(x, \Sigma_{\xi\xi} x) = \mathrm{E}(\xi, x)^2 \geqslant 0$. Из данного соотношения вытекает, что если $x \in \mathcal{N}(\Sigma_{\xi\xi})$, то $E(\xi, x)^2 = 0$, что вкупе с равенством $\mathrm{E}(\xi, x) = (\mathrm{E}\xi, x) = 0$ даёт $\mathrm{D}(\xi, x) = 0$. В силу неравенства Чебышёва

$$P\big(|(\xi, x)| > \varepsilon\big) \leqslant \frac{\mathrm{D}(\xi, x)}{\varepsilon^2} = 0$$

для любого $\varepsilon > 0$, поэтому $(\xi, x) = 0$ с вероятностью единица, если $x \in \mathcal{N}(\Sigma_{\xi\xi})$.

[3)] Мы вновь не вводим, если в этом нет необходимости, специальных символов для обозначения нулевых операторов, если они действуют в разных пространствах, и пишем $\Sigma_{\xi\eta} = 0$ и $\Sigma_{\eta\xi} = 0$.

Условие $\mathrm{E}\xi = 0$ и $\mathrm{E}\eta = 0$ можно заменить условием существования математических ожиданий. Тогда ковариационные операторы задаются равенствами

$$\begin{aligned}
(x, \Sigma_{\xi\xi} z) &= \mathrm{E}(\xi - \mathrm{E}\xi, x)(\xi - \mathrm{E}\xi, z), \qquad && x, z \in \mathcal{R}, \\
(y, \Sigma_{\xi\eta} x) &= \mathrm{E}(\xi - \mathrm{E}\xi, x)(\eta - \mathrm{E}\eta, y), \qquad && x \in \mathcal{R}, \quad y \in \widetilde{\mathcal{R}}. \\
(x, \Sigma_{\eta\xi} y) &= \mathrm{E}(\xi - \mathrm{E}\xi, x)(\eta - \mathrm{E}\eta, y), \qquad && x \in \mathcal{R}, \quad y \in \widetilde{\mathcal{R}}.
\end{aligned}$$

При этом соотношения $\Sigma_{\eta\xi} = \Sigma^*_{\xi\eta}$, $\Sigma_{\xi\xi} = \Sigma^*_{\xi}$ и $\Sigma_{\xi\xi} \geqslant 0$, разумеется, сохраняют свою силу, а условие $x \in \mathcal{N}(\Sigma_{\xi\xi})$ влечёт, что с вероятностью единица $(\xi, x) = (E\xi, x)$, т. е. (ξ, x) — опять же не случайная, а детерминированная величина.

Для упрощения формул (без потери общности, поскольку всегда можно совершить замену $\xi \mapsto \xi - \mathrm{E}\xi$) будем считать, что в последующих свойствах ковариационного оператора случайные элементы имеют нулевые математические ожидания.

1. Если $R \in \mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, то ковариационный оператор случайного элемента $R\xi$ есть $R\Sigma_{\xi\xi}R^*$, поскольку для любых $y, u \in \widetilde{\mathcal{R}}$

$$(\Sigma_{R\xi} y, u) = \mathrm{E}(R\xi, y)(R\xi, u) = \mathrm{E}(\xi, R^* y)(\xi, R^* u) = (\Sigma_{\xi\xi} R^* y, R^* u) = (R\Sigma_{\xi\xi} R^* y, u).$$

В силу произвольности u это означает, что $\Sigma_{R\xi} y = R\Sigma_{\xi\xi} R^* y$ для любого $y \in \widetilde{\mathcal{R}}$, что в свою очередь эквивалентно операторному равенству $\Sigma_{R\xi} = R\Sigma_{\xi\xi} R^*$.

2. Пусть $\xi_{1,2}$ — некоррелированные случайные элементы в $\mathcal{R}$. Тогда

$$\Sigma_{\xi_1+\xi_2, \xi_1+\xi_2} = \Sigma_{\xi_1} + \Sigma_{\xi_2}.$$

В самом деле, положим для краткости $\xi = \xi_1 + \xi_2$ и $\Sigma_{ij} = \Sigma_{\xi_i \xi_j}$, $i, j = 1, 2$, и, учитывая, что $\Sigma_{12} = 0$, $\Sigma_{21} = \Sigma^*_{12} = 0$, запишем цепочку равенств

$$\begin{aligned}
(\Sigma_{\xi_1+\xi_2} x, z) &= \mathrm{E}(\xi_1 + \xi_2, x)(\xi_1 + \xi_2, z) = \\
&= \mathrm{E}(\xi_1, x)(\xi_1, z) + \mathrm{E}(\xi_1, x)(\xi_2, z) + \mathrm{E}(\xi_2, x)(\xi_1, z) + \mathrm{E}(\xi_1, z)(\xi_1, z) = \\
&= (\Sigma_{11} x, z) + (\Sigma_{12} x, z) + (\Sigma_{21} z, x) + (\Sigma_{22} x, z) = \big((\Sigma_{11} + \Sigma_{22}) x, z\big)
\end{aligned}$$

для любых $x, z \in \mathcal{R}$. Отсюда $\Sigma_{\xi_1+\xi_2} x = \Sigma_{\xi_1\xi_1} x + \Sigma_{\xi_2\xi_2} x$ для любого $x \in \mathcal{R}$ и, следовательно, $\Sigma_{\xi_1+\xi_2} = \Sigma_{\xi_1} + \Sigma_{\xi_2}$.

3. Если ξ и η — гильбертовы случайные элементы и $R \in \mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, то имеют место равенства

$$\mathrm{E}\|\xi\|^2 = \mathrm{Tr}\,\Sigma_{\xi\xi}, \qquad \mathrm{E}(R\xi, \eta) = \mathrm{Tr}\,\Sigma_{\xi\eta} R^*. \tag{1.8}$$

Для доказательства заметим, что для любых ОНБ $\{e_i\}$ и $\{\tilde{e}_j\}$ пространств $\mathcal{R}$ и $\widetilde{\mathcal{R}}$ соответственно имеем

$$\begin{aligned}
\mathrm{E}\,\|\xi\|^2 &= \mathrm{E}\sum_{i=1}^{\infty}(\xi, e_i)^2 = \sum_{i=1}^{\infty}\mathrm{E}\,(\xi, e_i)^2 = \sum_{i=1}^{\infty}(\Sigma_{\xi\xi} e_i, e_i) = \mathrm{Tr}\,\Sigma_{\xi\xi},\\
\mathrm{E}\,(R\xi, \eta) &= \mathrm{E}\sum_{j=1}^{\infty}(R\xi, \tilde{e}_j)(\eta, \tilde{e}_j) = \sum_{j=1}^{\infty}\mathrm{E}\,(R\xi, \tilde{e}_j)(\eta, \tilde{e}_j) = \\
&= \sum_{j=1}^{\infty}\mathrm{E}\,(\xi, R^*\tilde{e}_j)(\eta, \tilde{e}_j) = \sum_{j=1}^{\infty}(\Sigma_{\xi\eta} R^*\tilde{e}_j, \tilde{e}_j) = \mathrm{Tr}\,\Sigma_{\xi\eta}R^*.
\end{aligned} \tag{1.9}$$

О законности внесения математического ожидания под знак бесконечной суммы в первом равенстве мы уже говорили (см. формулу (1.1)), однако подобная замена во второй цепочке равенств требует обоснования. Для этого заметим, что для любого $n = 1, 2, \ldots$ в силу неравенств Коши–Буняковского и Бесселя

$$\left|\sum_{j=1}^{n}(R\xi, \tilde{e}_j)(\eta, \tilde{e}_j)\right| \leqslant \sum_{j=1}^{n}|(R\xi, \tilde{e}_j)(\eta, \tilde{e}_j)| \leqslant \sqrt{\sum_{j=1}^{n}(R\xi, \tilde{e}_j)^2}\sqrt{\sum_{j=1}^{n}(\eta, \tilde{e}_j)^2} = \|R\xi\|\cdot\|\eta\|.$$

Если $\mathrm{E}\,\|R\xi\|^2 < \infty$, $\mathrm{E}\,\|\eta\|^2 < \infty$, то

$$E\big(\|R\xi\|\cdot\|\eta\|\big) \leqslant \sqrt{\mathrm{E}\,\|R\xi\|^2}\sqrt{\mathrm{E}\,\|\eta\|^2} < \infty, \tag{1.10}$$

где мы воспользовались вновь неравенством Коши–Буняковского, но теперь для математических ожиданий случайных величин $\|R\xi\|$ и $\|\eta\|$. Таким образом, можно применить теорему A.2, приняв во внимание, что

$$\sum_{j=1}^{n}(R\xi, \tilde{e}_j)(\eta, \tilde{e}_j) \underset{n\to\infty}{\longrightarrow} \sum_{j=1}^{\infty}(R\xi, \tilde{e}_j)(\eta, \tilde{e}_j), \qquad \left|\sum_{j=1}^{n}(R\xi, \tilde{e}_j)(\eta, \tilde{e}_j)\right| \leqslant \|R\xi\|\cdot\|\eta\|.$$

Свойство доказано.

Понятно, что соотношения (1.9) не зависят от выбора ОНБ $\{e_i\}$ и $\{\tilde{e}_j\}$ в пространствах $\mathcal{R}$ и $\widetilde{\mathcal{R}}$.Таким образом, для любого элемента $x \in \mathcal{R}$, нормированного условием $\|x\| = 1$ мы можем, положив $e_1 = x$, написать

$$(\Sigma_{\xi\xi} x, x) = (\Sigma_{\xi\xi} e_1, e_1) \leqslant \sum_{i=1}^{\infty}(\Sigma_{\xi\xi} e_i, e_i) = \mathrm{Tr}\,\Sigma_{\xi\xi},$$

где мы воспользовались тем, что $\Sigma_{\xi\xi} \geqslant 0$. Отсюда (см. теорему B.22 приложения B)

$$\|\Sigma_{\xi\xi}\| = \sup_{\|x\|=1}(\Sigma_{\xi\xi} x, x) \leqslant \mathrm{Tr}\,\Sigma_{\xi\xi} < \infty,$$

т. е. оператор $\Sigma_{\xi\xi} \geqslant 0$ ограничен. Следовательно, существует квадратный корень из этого оператора (см. теорему B.11) — ограниченный самосопряжённый неотрицательно определенный оператор $\Sigma_\xi^{1/2}$ такой, что $\Sigma_\xi^{1/2}\Sigma_\xi^{1/2} = \Sigma_{\xi\xi}$. При этом

$$\mathrm{E}\,\|\xi\|^2 = \mathrm{Tr}\,\Sigma_{\xi\xi} = \mathrm{Tr}\,\Sigma_\xi^{1/2}\Sigma_\xi^{1/2} = \mathrm{Tr}\,\Sigma_\xi^{1/2}\big(\Sigma_\xi^{1/2}\big)^* = \|\Sigma_\xi^{1/2}\|_2^2. \tag{1.11}$$

Итак, мы имеем ещё одно свойство.

4. Случайный элемент ξ является гильбертовым случайным элементом тогда и только тогда, когда $\Sigma_\xi^{1/2}$ принадлежит пространству операторов Гильберта–Шмидта, и в этом случае $\mathrm{E}\,\|\xi\|^2 = \|\Sigma_\xi^{1/2}\|_2^2$.

Если $R \in \mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ и $\Sigma_\xi^{1/2} \in \mathbb{H}(\mathcal{R} \mapsto \mathcal{R})$, то

$$R\Sigma_{\xi\xi}R^* = R\Sigma_\xi^{1/2}\Sigma_\xi^{1/2}R^* \in \mathbb{H}(\widetilde{\mathcal{R}} \mapsto \widetilde{\mathcal{R}})$$

как произведение ограниченных операторов и операторов Гильберта–Шмидта. Таким образом, $R\xi$ — тоже гильбертов случайный элемент. При этом аналогично (1.11)

$$\mathrm{E}\,\|R\xi\|^2 = \mathrm{Tr}\,R\Sigma_\xi^{1/2}\big(R\Sigma_\xi^{1/2}\big)^* = \|R\Sigma_\xi^{1/2}\|_2^2. \tag{1.12}$$

Наконец, если $R\xi$ и η — гильбертовы случайные элементы, то с учётом неравенства (1.10) и очевидного неравенства $|\mathrm{E}\alpha| \leqslant \mathrm{E}|\alpha|$ для математического ожидания случайной величины можно записать

$$|\mathrm{E}(R\xi,\eta)| \leqslant \mathrm{E}|(R\xi,\eta)| \leqslant \mathrm{E}\big(\|R\xi\|\cdot\|\eta\|\big) \leqslant \sqrt{\mathrm{E}\|R\xi\|^2}\sqrt{\mathrm{E}\|\eta\|^2}.$$

таким образом, применяя (1.12) и аналог равенства (1.11) для случайного элемента η, имеем следующее свойство.

5. Справедливо неравенство

$$|\mathrm{E}(R\xi,\eta)| \leqslant \|R\Sigma_\xi^{1/2}\|_2 \cdot \|\Sigma_\eta^{1/2}\|_2. \tag{1.13}$$

6. Если потребовать включение $R, \Sigma_{\xi\eta} \in \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, то в силу второго из равенств (1.8)

$$\mathrm{E}(R\xi,\eta) = \mathrm{Tr}\,\Sigma_{\xi\eta}R^* = (\Sigma_{\xi\eta}, R)_2. \tag{1.14}$$

1.3. Оценивание случайных элементов. Пусть доступна наблюдению реализация случайного элемента ξ со значениями в $\mathcal{R}$, а η — случайный элемент в $\widetilde{\mathcal{R}}$, коррелированный с ξ. Поставим задачу найти наилучшую в среднем квадратичном смысле оценку для η среди всех оценок, заданных как линейное преобразование элемента ξ. Другими словами, требуется найти линейный оператор R из $\mathcal{R}$ в $\widetilde{\mathcal{R}}$, доставляющий $\min \mathrm{E}\|R\xi - \eta\|^2$.

Уточним модель наблюдения. Предположим, что $\mathrm{E}\xi = 0$, $\mathrm{E}\eta = 0$ и что существуют ковариационные операторы $\Sigma_{\xi\xi} = \Sigma$, $\Sigma_{\eta\eta}$ и взаимный ковариационный оператор $\Sigma_{\xi\eta}$. Потребуем, чтобы два последних оператора были операторами Гильберта–Шмидта,

$$\Sigma_{\eta\eta} \in \mathbb{H}(\widetilde{\mathcal{R}} \mapsto \widetilde{\mathcal{R}}), \qquad \Sigma_{\xi\eta} \in \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}}).$$

Для оператора Σ потребуем включение $\Sigma \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$, чтобы существовал $\Sigma^{1/2}$. Кроме того, вспомним, что если $x \in \mathcal{N}(\Sigma)$, то $(\xi, x) = 0$ с вероятностью единица, поэтому составляющая случайного элемента ξ, лежащая в $\mathcal{N}(\Sigma)$, не несёт информации об η и может быть исключена из процедуры оценивания. Это означает, что мы можем сузить пространство $\mathcal{R}$ до $\mathcal{N}^{\perp}(\Sigma)$, тогда Σ станет невырожденным оператором. При этом после такого сужения

$$\mathcal{N}(\Sigma^{1/2}) = \mathcal{N}\big((\Sigma^{1/2})^*\Sigma^{1/2}\big) = \mathcal{N}\big(\Sigma^{1/2}\Sigma^{1/2}\big) = \mathcal{N}(\Sigma) = \{0\},$$

поэтому существует оператор $(\Sigma^{1/2})^{-1}$, который мы будем обозначать как $\Sigma^{-1/2}$. Заметим, что если $\Sigma \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$, то $\Sigma^{1/2} \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$ и $\Sigma^{-1}\Sigma = \Sigma^{-1/2}\Sigma^{1/2} = I$, но операторы $\Sigma\Sigma^{-1}$ и $\Sigma^{1/2}\Sigma^{-1/2}$, вообще говоря, отличны от единичного, т. к. определены на линейных многообразиях $\mathcal{R}(\Sigma)$ и $\mathcal{R}(\Sigma^{1/2})$. Эти многообразия плотны в $\mathcal{R}$ в силу $\mathcal{R}^{\perp}(\Sigma^{1/2}) = \mathcal{N}(\Sigma^{1/2}) = \{0\}$ и $\mathcal{R}^{\perp}(\Sigma^{1/2}) = \mathcal{N}(\Sigma^{1/2}) = \{0\}$, но не обязательно совпадают с $\mathcal{R}$ (многообразия могут быть незамкнуты, и тогда соответствующие обратные операторы неограничены по теореме о замкнутом графике).

Итак, будем считать, что

$$\mathcal{N}(\Sigma^{1/2}) = \{0\}, \tag{1.15}$$

и ведем следующее множество линейных операторов:

$$\mathbb{H}_{\Sigma^{1/2}} = \big\{R \in \mathbb{L}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})\colon\ \mathcal{D}(R) = \mathcal{R}(\Sigma^{1/2}),\ \overline{R\Sigma^{1/2}} \in \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})\big\}. \tag{1.16}$$

Определим на $\mathbb{H}_{\Sigma^{1/2}}$ стандартные линейные операции и зададим норму:

$$\|R\|_{\Sigma^{1/2}} = \|R\Sigma^{1/2}\|_2, \qquad R \in \mathbb{H}_{\Sigma^{1/2}}.$$

Из всех аксиом нормы отметим, что $\|R\|_{\Sigma^{1/2}} = \|R\Sigma^{1/2}\|_2 = 0$ влечёт $R\Sigma^{1/2} = 0$, т. е. $Ry = 0$ для любого $y \in \mathcal{R}(\Sigma^{1/2})$ или, другими словами, $R = 0$ как оператор из пространства $\mathbb{H}_{\Sigma^{1/2}}$. При этом как и любой другой ограниченный оператор, заданный на незамкнутом всюду плотном линейном многообразии, оператор $R = 0$ на $\mathcal{R}(\Sigma^{1/2})$ можно доопределить (по непрерывности) как нулевой оператор во всём $\mathcal{R} = \overline{\mathcal{R}(\Sigma^{1/2})}$.

Введённая норма порождена скалярным произведением, заданным формулой

$$(R_1, R_2)_{\Sigma^{1/2}} = (R_1\Sigma^{1/2}, R_2\Sigma^{1/2})_2, \quad R_{1,2} \in \mathbb{H}_{\Sigma^{1/2}}.$$

Теорема 1.

Пространство $\mathbb{H}_{\Sigma^{1/2}}$ полно по норме $\|\cdot\|_{\Sigma^{1/2}}$. Пространство $\mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ есть всюду плотное (по норме пространства $\mathbb{H}_{\Sigma^{1/2}}$) подмножество в $\mathbb{H}_{\Sigma^{1/2}}$.

Доказательство. Пусть последовательность $\{R_n\} \subset \mathbb{H}_{\Sigma^{1/2}}$ фундаментальна, т. е. $\|R_n - R_m\|_{\Sigma^{1/2}} = \|R_n\Sigma^{1/2} - R_m\Sigma^{1/2}\|_2 \to 0$ при $n, m \to \infty$. В силу полноты пространства $\mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ последовательность $\{R_n\Sigma^{1/2}\} \subset \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ сходится к некоторому оператору $Q \in \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$.

Зададим на $\mathcal{R}(\Sigma^{1/2})$ оператор R. Пусть $z = \Sigma^{1/2}x \in \mathcal{R}(\Sigma^{1/2})$ при некотором $x \in \mathcal{R}$, причём в силу обратимости $\Sigma^{1/2}$ такой элемент x единствен. Положим $Rz = R\Sigma^{1/2}x = Qx$ для любого $x \in \widetilde{\mathcal{R}}$, тогда $R\Sigma^{1/2} = Q \in \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, и $R \in \mathbb{H}_{\Sigma^{1/2}}$ есть предел последовательности $\{R_n\}$ по норме пространства $\mathbb{H}_{\Sigma^{1/2}}$:

$$\|R - R_m\|_{\Sigma^{1/2}} = \|R\Sigma^{1/2} - R_m\Sigma^{1/2}\|_2 = \|Q - R_m\Sigma^{1/2}\|_2 \xrightarrow[n\to\infty]{} 0.$$

Полнота пространства $\mathbb{H}_{\Sigma^{1/2}}$ доказана.

Если $R \in \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, то $R\Sigma^{1/2} \in \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ как произведение ограниченного оператора и оператора Гильберта–Шмидта. Таким образом, $\mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}}) \subset \mathbb{H}_{\Sigma^{1/2}}$. Возьмем произвольный оператор $R \in \mathbb{H}_{\Sigma^{1/2}}$. Оператор $R\Sigma^{1/2}$ принадлежит пространству $\mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, поэтому имеет сингулярные базисы $\{e_i\}$ и $\{\tilde{e}_i\}$ (см. приложение C):

$$R\Sigma^{1/2}e_i = \beta_i\tilde{e}_i, \qquad R\Sigma^{1/2}x = \sum_{i=1}^{\infty} \beta_i(x, e_i)\tilde{e}_i, \qquad \|R\Sigma^{1/2}\|_2 = \sum_{i=1}^{\infty} \beta_i^2 < \infty.$$

Положим для $z = \Sigma^{1/2}x \in \mathcal{R}(\Sigma^{1/2})$

$$R_n z = R_n\Sigma^{1/2}x = \sum_{i=1}^{n} \beta_i(x, e_i)\tilde{e}_i.$$

Тогда $\mathcal{R}(R_n) \subset \mathcal{L}(\tilde{e}_1, \ldots, \tilde{e}_n)$ и $\operatorname{rank} R_n \leqslant n < \infty$. Таким образом, оператор R_n является ограниченным на (всюду плотном) линейном многообразии $\mathcal{R}(\Sigma^{1/2})$ оператором как оператор конечного ранга. Его доопределение по непрерывности на всё пространство $\mathcal{R} = \overline{\mathcal{R}(\Sigma^{1/2})}$ принадлежит $\mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ опять же как оператор конечного ранга. С другой стороны, по определению

$$R_n\Sigma^{1/2}e_k = \begin{cases} \beta_k\tilde{e}_k = Re_k, & k \leqslant n, \\ 0, & k > n, \end{cases}$$

поэтому при $n \to \infty$

$$\|R\Sigma^{1/2} - R_n\Sigma^{1/2}\|_2^2 = \sum_{i=1}^{\infty} \|(R\Sigma^{1/2} - R_n\Sigma^{1/2})e_i\|^2 = \sum_{i=n+1}^{\infty} \|R\Sigma^{1/2}e_i\|^2 = \sum_{i=n+1}^{\infty} \beta_i^2 \to 0.$$

Таким образом $\|R - R_n\|_{\Sigma^{1/2}} \to 0$, и это доказывает, что любой оператор из $\mathbb{H}_{\Sigma^{1/2}}$ есть предел последовательности операторов из $\mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ по норме $\|\cdot\|_{\Sigma^{1/2}}$. Другими словами, $\overline{\mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})} = \mathbb{H}_{\Sigma^{1/2}}$. Теорема доказана.

Вернемся к задаче оценивания.

Теорема 2.
Пусть ξ — случайный элемент со значениями в $\mathcal{R}$, а η — случайный элемент со значениями в $\widetilde{\mathcal{R}}$, коррелированный с ξ. Предположим, что $\mathrm{E}\xi = 0$, $\mathrm{E}\eta = 0$ и существуют ковариационные операторы $\Sigma_{\xi\xi} = \Sigma$, $\Sigma_{\eta\eta}$ и взаимный ковариационный оператор $\Sigma_{\xi\eta}$, причём

$$\Sigma_{\xi\xi} \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R}), \qquad \Sigma_{\eta\eta} \in \mathbb{H}(\widetilde{\mathcal{R}} \mapsto \widetilde{\mathcal{R}}), \qquad \Sigma_{\xi\eta} \in \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}}).$$

Пусть

$$\mathbb{H}_{\Sigma^{1/2}} = \left\{R \in \mathbb{L}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})\colon\ \mathcal{D}(R) = \mathcal{R}(\Sigma^{1/2}),\ R\Sigma^{1/2} \in \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})\right\}.$$

Тогда

$$\min_{R\in\mathbb{H}_{\Sigma^{1/2}}} \mathrm{E}\|R\xi - \eta\|^2 = \mathrm{E}\|R_\circ\xi - \eta\|^2 = \|\Sigma_\eta^{1/2}\|_2^2 - \|\overline{\Sigma_{\xi\eta}\Sigma^{-1/2}}\|_2^2, \tag{1.17}$$

где

$$R_\circ = \overline{\Sigma_{\xi\eta}\Sigma^{-1/2}} \cdot \Sigma^{-1/2}.$$

Доказательство. Пусть $R \in \mathbb{H}_{\Sigma^{1/2}}$. Запишем среднеквадратичную погрешность, пользуясь свойствами ковариационных операторов:

$$\begin{aligned}\mathrm{E}\|R\xi - \eta\|^2 &= \mathrm{E}\|R\xi\|^2 - 2\mathrm{E}(R\xi, \eta) + \mathrm{E}\|\eta\|^2 = \|R\Sigma^{1/2}\|_2^2 - 2\mathrm{E}(R\xi, \eta) + \|\Sigma_\eta^{1/2}\|_2^2 = \\ &= \|R\|_{\Sigma^{1/2}}^2 - 2\mathrm{E}(R\xi, \eta) + \|\Sigma_\eta^{1/2}\|_2^2.\end{aligned} \tag{1.18}$$

Теперь представим второе слагаемое в правой части равенства (1.18) в удобной для решения экстремальной задачи виде: мы покажем, что для некоторого $R_\circ \in \mathbb{H}_{\Sigma^{1/2}}$

$$\mathrm{E}(R\xi, \eta) = (R, R_\circ)_{\Sigma^{1/2}}, \qquad R \in \mathbb{H}_{\Sigma^{1/2}}. \tag{1.19}$$

Сначала сузим пространство $\mathbb{H}_{\Sigma^{1/2}}$ до всюду плотного линейного многообразия $\mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$. Для каждого $R \in \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ положим $f(R) = \mathrm{E}(R\xi, \eta) = (R, \Sigma_{\xi\eta})_2$

(см. равенство (1.14)). Понятно, что таким образом мы получим $f(\,\cdot\,)$ — линейный функционал, заданный на всюду плотном подмножестве $\mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ гильбертова пространства $\mathbb{H}_{\Sigma^{1/2}}$. При этом в силу (1.13) $f(\,\cdot\,)$ ограничен как функционал над $\mathbb{H}_{\Sigma^{1/2}}$:

$$|f(R)| = |\mathrm{E}(R\xi, \eta)| \leqslant \|R\Sigma^{1/2}\|_2 \|\Sigma_\eta^{1/2}\|_2 = \|R\|_{\Sigma^{1/2}} \|\Sigma_\eta^{1/2}\|_2,$$

следовательно, $\|f\| \leqslant \|\Sigma_\eta^{1/2}\|_2$. Тогда его можно доопределить по непрерывности на всё пространство $\mathbb{H}_{\Sigma^{1/2}} = \overline{\mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})}$ как функционал $\bar{f}$, причём $\|f\| = \|\bar{f}\|$. По тереме Рисса об общем виде линейного функционала найдется единственный элемент $R_\circ \in \mathbb{H}_{\Sigma^{1/2}}$ такой, что

$$\bar{f}(R) = (R, R_\circ)_{\Sigma^{1/2}} = \mathrm{E}(R\xi, \eta) = (R, R_\circ)_{\Sigma^{1/2}}, \qquad R \in \mathbb{H}_{\Sigma^{1/2}}.$$

Теперь мы без труда можем решить задачу на минимум. Подставляем последнее равенство в правую часть равенства (1.18):

$$\mathrm{E}\|R\xi - \eta\|^2 = \|R\|^2_{\Sigma^{1/2}} - 2(R, R_\circ)_{\Sigma^{1/2}} + \|\Sigma_\eta^{1/2}\|_2^2 = \|R - R_\circ\|^2_{\Sigma^{1/2}} - \|R_\circ\|^2_{\Sigma^{1/2}} + \|\Sigma_\eta^{1/2}\|_2^2.$$

Отсюда сразу вытекает, что для любого $R \in \mathbb{H}_{\Sigma^{1/2}}$

$$\mathrm{E}\|R\xi - \eta\|^2 \geqslant \|\Sigma_\eta^{1/2}\|_2^2 - \|R_\circ\|^2_{\Sigma^{1/2}}.$$

причем равенство возможно тогда и только тогда, когда $\|R - R_\circ\|^2_{\Sigma^{1/2}} = 0$, т. е. выполнено равенство $R\Sigma^{1/2} = R_\circ\Sigma^{1/2}$. Таким образом,

$$\min_{R \in \mathbb{H}_{\Sigma^{1/2}}} \mathrm{E}\|R\xi - \eta\|^2 = \mathrm{E}\|R_\circ\xi - \eta\|^2 = \|\Sigma_\eta^{1/2}\|_2^2 - \|R_\circ\|^2_{\Sigma^{1/2}}. \tag{1.20}$$

Осталось найти $R_\circ$. По условию $\mathrm{E}(R\xi, \eta) = (R, \Sigma_{\xi\eta})_2 = (R, R_\circ)_{\Sigma^{1/2}}$ для любого $R \in \mathbb{H}_{\Sigma^{1/2}}$. В результате

$$(\Sigma_{\xi\eta}, R)_2 = (R_\circ\Sigma^{1/2}, R\Sigma^{1/2})_2 = \operatorname{Tr} R_\circ\Sigma^{1/2}(R\Sigma^{1/2})^* = \operatorname{Tr} R_\circ\Sigma R^* = (R_\circ\Sigma, R)_2$$

для любого $R \in \mathbb{H}_{\Sigma^{1/2}}$. Следовательно, $\Sigma_{\xi\eta} = R_\circ\Sigma$. Перепишем последнее равенство как $R_\circ\Sigma^{1/2} \cdot \Sigma^{1/2} = \Sigma_{\xi\eta}$. Отсюда получим

$$R_\circ\Sigma^{1/2}x = \Sigma_{\xi\eta}\Sigma^{-1/2}x, \qquad x \in \mathcal{D}(\Sigma^{-1/2}) = \mathcal{R}(\Sigma^{-1/2}).$$

По условию $R_\circ \in \mathbb{H}_{\Sigma^{1/2}}$, т. е. $R_\circ\Sigma^{1/2} \in \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ и, следовательно,

$$\|\Sigma_{\xi\eta}\Sigma^{-1/2}\| = \sup_{\substack{x \in \mathcal{R}(\Sigma^{-1/2}),\\ \|x\| \leqslant 1}} \|\Sigma_{\xi\eta}\Sigma^{-1/2}x\| = \sup_{\substack{x \in \mathcal{R}(\Sigma^{-1/2}),\\ \|x\| \leqslant 1}} \|R_\circ\Sigma^{1/2}x\| < \infty.$$

Замкнём ограниченный на $\mathcal{R}(\Sigma^{-1/2})$ оператор $\Sigma_{\xi\eta}\Sigma^{-1/2}$ по непрерывности, получим, что оператор $\overline{\Sigma_{\xi\eta}\Sigma^{-1/2}} = R_\circ\Sigma^{1/2} \in \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$. Таким образом,

$$R_\circ = \overline{\Sigma_{\xi\eta}\Sigma^{-1/2}} \cdot \Sigma^{-1/2}. \tag{1.21}$$

Тогда

$$\|R_\circ\|^2_{\Sigma^{1/2}} = \|R_\circ\Sigma^{1/2}\|^2_2 = \|\overline{\Sigma_{\xi\eta}\Sigma^{-1/2}}\|^2_2, \tag{1.22}$$

причём, выбрав для расчёта последней нормы ОНБ во всюду плотном линейном многообразии $\mathcal{R}(\Sigma^{1/2}) = \mathcal{D}(\Sigma^{-1/2})$, мы можем не писать в ней символ замыкания.

Подставляя $R_\circ$ из (1.21) и выражение для $\|R_\circ\|^2_{\Sigma^{1/2}}$ из (1.22) в (1.20), получаем (1.17). Теорема доказана.

1.4. Задача редукции измерения. Пусть

$$\begin{aligned} &\xi = A\phi + \nu, \qquad A \in \mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}}), \\ &\phi \text{ — случайный элемент в } \mathcal{R}, \qquad \mathrm{E}\phi = f_0, \quad \Sigma_\phi = F, \\ &\nu \text{ — случайный элемент в } \widetilde{\mathcal{R}}, \qquad \mathrm{E}\nu = 0, \quad \Sigma_\nu = \Sigma. \end{aligned} \tag{1.23}$$

Будем считать, что случайные элементы ϕ и ν некоррелированы, $\Sigma_{\phi\nu} = 0$. Будем говорить, что условия (1.23) задают *линейную модель* $[A, f_0, F, \Sigma]$ *измерения* случайного элемента ϕ. Действие оператора A моделирует воздействие измерительного прибора на измеряемый сигнал ϕ, случайный элемент ν — это модель случайной погрешности (шума) измерения, условие $\mathrm{E}\nu = 0$ отвечает отсутствию систематической погрешности.

Пусть задан U — линейный ограниченный всюду определённый ограниченный оператор из $\mathcal{R}$ в некоторое гильбертово пространство $\widehat{\mathcal{R}}$ и $\eta = U\phi$. Тогда в силу общих свойств математического ожидания и ковариационного оператора мы имеем $\mathrm{E}_\eta = Uf_0$ и $\Sigma_{\eta\eta} = UFU^*$. Кроме того, ковариационный оператор $S = \Sigma_{\xi\xi}$ при условии некоррелированности ϕ и ν и, следовательно, некоррелированности $A\phi$ и ν равен $S = \Sigma_{A\phi} + \Sigma_\nu = AFA^* + \Sigma$.

Поставим следующую вариационную задачу: требуется найти линейный оператор R, действующий из $\widetilde{\mathcal{R}}$ в $\widehat{\mathcal{R}}$, и (неслучайный) элемент $r \in \widehat{\mathcal{R}}$, которые доставляют

$$\min_R \min_{r\in\widehat{\mathcal{R}}} \mathrm{E}\,\|R\xi + r - U\phi\,\|^2, \tag{1.24}$$

где математическое ожидание вычисляется по совместному распределению случайных элементов ϕ и ν (множество операторов R, по которому вычисляется минимум, будет указано ниже, пока потребуем только $\mathcal{D}(R) = \widetilde{\mathcal{R}}$). «Центрируем» входящие в (1.24) случайные элементы, т. е. вычтем их математические ожидания:

$$\phi^\circ = \phi - \mathrm{E}\phi = \phi - f_0,$$
$$\xi^\circ = \xi - \mathrm{E}\xi = \xi - (\mathrm{E}A\phi + \mathrm{E}\nu) = \xi - Af_0 = A(\phi - f_0) + \nu = A\phi^\circ + \nu,$$
$$\eta^\circ = \eta - \mathrm{E}\eta = U\phi - Uf_0 = U(\phi - f_0) = U\phi^\circ.$$

Тогда погрешность в (1.24) записывается как

$$\mathrm{E}\|R\xi + r - U\phi\|^2 = \mathrm{E}\|R(\xi^\circ + Af_0) + r - (U\phi^\circ + Uf_0)\|^2 = \mathrm{E}\|R\xi^\circ + r^\circ - U\phi^\circ\|^2,$$

где $r^\circ = r + RAf_0 - Uf_0$. При этом, очевидно,

$$\begin{aligned} &\mathrm{E}\xi^\circ = 0, \qquad \Sigma_{\xi^\circ\xi^\circ} = \Sigma_{\xi\xi} = S = AFA^* + \Sigma, \\ &\mathrm{E}\eta^\circ = 0, \qquad \Sigma_{\eta^\circ\eta^\circ} = \Sigma_{\eta\eta} = UFU^*. \end{aligned} \tag{1.25}$$

Используя полученные соотношения, преобразуем погрешность:

$$\begin{aligned} \mathrm{E}\|R\xi + r - U\phi\|^2 &= \mathrm{E}\|R\xi^\circ + r^\circ - U\phi^\circ\|^2 = \\ &= \mathrm{E}\|R\xi^\circ - U\phi^\circ\|^2 - 2\mathrm{E}(R\xi^\circ - U\phi^\circ, r^\circ) + \|r^\circ\|^2 = \\ &= \mathrm{E}\|R\xi^\circ - U\phi^\circ\|^2 - 2(\mathrm{E}(R\xi^\circ - U\phi^\circ), r^\circ) + \|r^\circ\|^2 = \\ &= \mathrm{E}\|R\xi^\circ - U\phi^\circ\|^2 + \|r^\circ\|^2 \geqslant \mathrm{E}\|R\xi^\circ - U\phi^\circ\|^2, \end{aligned}$$

где мы воспользовались тем, что $\mathrm{E}(R\xi^\circ - U\phi^\circ) = \mathrm{E}R\xi^\circ - \mathrm{E}\,U\phi^\circ = 0$. Отсюда сразу вытекает, что

$$\min_{r\in\widehat{\mathcal{R}}} \mathrm{E}\|R\xi + r - U\phi\|^2 = \min_{r^\circ\in\widehat{\mathcal{R}}} \mathrm{E}\|R\xi^\circ + r^\circ - U\phi^\circ\|^2 = \mathrm{E}\|R\xi^\circ - U\phi^\circ\|^2 \tag{1.26}$$

и достигается на единственном элементе $r^\circ = 0$, т. е. при $r = -RAf_0 + Uf_0$. Заметим, что при этом

$$\mathrm{E}(R\xi + r) = \mathrm{E}R\xi + r = RAf_0 - RAf_0 + Uf_0 = Uf_0 = \mathrm{E}\,U\phi = \mathrm{E}\eta,$$

другими словами, $R\xi + r$ — несмещённая оценка случайного элемента η.

Теперь задача (1.24) свелась к задаче, рассмотренной в теореме 2: требуется найти $\min \mathrm{E}\|R\xi^\circ - \eta^\circ\|$, где ковариационные операторы случайных элементов ξ° и η° определены в (1.25). Для завершения постановки вариационной задачи и её решения нам осталось найти взаимный ковариационный оператор $\Sigma_{\xi^\circ\eta^\circ}$. Имеем для $y \in \widetilde{\mathcal{R}}$ и $z \in \widehat{\mathcal{R}}$

$$\begin{aligned} \mathrm{E}(\xi^\circ, y)(\eta^\circ, z) &= \mathrm{E}(A\phi^\circ + \nu, y)(U\phi^\circ, z) = \mathrm{E}(A\phi^\circ, y)(U\phi^\circ, z) + \mathrm{E}(\nu, y)(U\phi^\circ, z) = \\ &= \mathrm{E}(\phi^\circ, A^*y)(\phi^\circ, U^*z) + \mathrm{E}(\nu, y)\cdot\mathrm{E}(U\phi^\circ, z) = \\ &= (\Sigma_{\phi^\circ}A^*y, U^*z) = (\Sigma_\phi A^*y, U^*z) = (UFA^*y, z), \end{aligned}$$

где мы воспользовались некоррелированностью случайных элементов ν и $U\phi^\circ$ и равенствами $\mathrm{E}\nu = 0$ и $\mathrm{E}\,U\phi^\circ = 0$. Отсюда $\Sigma_{\xi^\circ\eta^\circ} = UFA^*$. Теперь мы готовы воспользоваться теоремой 2.

Пусть $\xi = A\phi + \nu$ — случайный элемент из (1.23) и $U\phi$ — линейное преобразование случайного элемента ϕ, $U \in \mathbb{B}(\mathcal{R} \mapsto \widehat{\mathcal{R}})$. Предположим, что ковариационные операторы $\Sigma_{\xi\xi} = S$, $\Sigma_{U\phi} = UFU^*$ и $\Sigma_{\xi\eta} = UFA^*$ удовлетворяют следующим включениям:

$$S \in \mathbb{B}(\widetilde{\mathcal{R}} \mapsto \widetilde{\mathcal{R}}), \qquad UFU^* \in \mathbb{H}(\widehat{\mathcal{R}} \mapsto \widehat{\mathcal{R}}), \qquad UFA^* \in \mathbb{H}(\widetilde{\mathcal{R}} \mapsto \widehat{\mathcal{R}}).$$

Пусть

$$\mathbb{H}_{S^{1/2}} = \bigl\{R \in \mathbb{L}(\widetilde{\mathcal{R}} \mapsto \widehat{\mathcal{R}})\colon \mathcal{D}(R) = \mathcal{R}(S^{1/2}),\ RS^{1/2} \in \mathbb{H}(\widetilde{\mathcal{R}} \mapsto \widehat{\mathcal{R}})\bigr\}.$$

Тогда

$$\min_{R\in\mathbb{H}_{S^{1/2}}} \min_{r\in\widehat{\mathcal{R}}} \mathrm{E}\,\|R\xi + r - U\phi\,\|^2 = \mathrm{E}\,\|R_0\xi - U\phi\,\|^2, \tag{1.27}$$

$$R_0 = \overline{UFA^*(AFA^* + \Sigma)^{-1/2}}\,(AFA^* + \Sigma)^{-1/2}. \tag{1.28}$$

Вариационная задача, сформулированная в (1.27), называется *задачей редукции измерения $\xi = A\phi + \nu$ к виду $\eta = U\phi$ в модели* $[A, f_0, F, \Sigma]$. Как мы уже отмечали, условие минимума по $r \in \widehat{\mathcal{R}}$ можно заменить $r = 0$ и условием несмещённости $\mathrm{E}(R\xi + r) = \mathrm{E}\,U\phi$.

По теореме 2 оператор R_0, заданный в (1.28), доставляет минимум. Он называется *оператором редукции.*

Погрешность редукции с учетом (1.17) и вида ковариационных операторов может быть записана как

$$\begin{aligned}\mathrm{E}\,\|R_0\xi - U\phi\,\|^2 &= \|\Sigma_\eta^{1/2}\|_2^2 - \|\overline{\Sigma_{\xi\eta}\Sigma^{-1/2}}\|_2 = \\ &= \|(UFU^*)^{1/2}\|_2^2 - \|\overline{UFA^*(AFA^* + \Sigma)^{-1/2}}\|_2 = \\ &= \operatorname{Tr} UFU^* - \operatorname{Tr} UFA^*(AFA^* + \Sigma)^{-1}(UFA^*)^*,\end{aligned}$$

где мы воспользовались определением $\|B\|_2 = \operatorname{Tr} BB^*$ и самосопряжённостью операторов $(UFU^*)^{1/2}$ и $(AFA^* + \Sigma)^{-1/2}$.

ГЛАВА 2
ПСЕВДООБРАЩЕНИЕ ОПЕРАТОРОВ

Идея псевдообращения широко используется в теории конечномерных линейных операторов для решения систем линейных уравнений, уравнений в линейных операторах и вариационных задач. Одним из основных свойств оператора A^{-}, псевдообратного линейному оператору A, является, например, равенство

$$\min_x \|Ax - y\| = \|AA^{-}y - y\|,$$

из которого следует, что элемент $A^{-}y$ гарантирует наилучшее приближение элемента y элементами вида Ax. Помимо этого, можно отметить, что псевдообратный оператор обладает значительно менее стеснёнными условиями существования по сравнению с обратным, но в случае, когда обратный оператор существует, совпадает с ним.

Мы введем псевдообратный оператор для любого плотно определённого замкнутого линейного оператора, который действует из бесконечномерного действительного гильбертова пространства $\mathcal{R}$ в бесконечномерное действительное гильбертово пространство $\widetilde{\mathcal{R}}$; исследуем его свойства и применения к решению вариационных задач и линейных уравнений.

2.1. Определение псевдообратного оператора. Пусть A — замкнутый, плотно определённый линейный оператор из $\mathcal{R}$ в $\widetilde{\mathcal{R}}$. Тогда $\mathcal{N}(A)$ и $\mathcal{N}(A^*)$ суть линейные подпространства в $\mathcal{R}$ и $\widetilde{\mathcal{R}}$, следовательно, справедливы разложения в прямые суммы подпространств

$$\mathcal{R} = \mathcal{N}(A) \oplus \mathcal{N}^{\perp}(A), \qquad \widetilde{\mathcal{R}} = \mathcal{N}(A^*) \oplus \mathcal{N}^{\perp}(A^*).$$

В силу равенств $\mathcal{N}^{\perp}(A) = \overline{\mathcal{R}(A^*)}$, $\mathcal{N}(A) = \mathcal{R}^{\perp}(A^*)$ и аналогичных равенств для $\mathcal{N}^{\perp}(A^*)$ и $\mathcal{N}(A^*)$ (см. приложение B, теорема B.7, формула (B.14)) данные разложения в прямые суммы подпространств могут быть записаны в разных формах:

$$\begin{gathered}\mathcal{N}(A) \oplus \mathcal{N}^{\perp}(A) = \mathcal{N}(A) \oplus \overline{\mathcal{R}(A^*)} = \mathcal{R}^{\perp}(A^*) \oplus \mathcal{N}^{\perp}(A) = \mathcal{R}^{\perp}(A^*) \oplus \overline{\mathcal{R}(A^*)},\\ \mathcal{N}(A^*) \oplus \mathcal{N}^{\perp}(A^*) = \mathcal{N}(A^*) \oplus \overline{\mathcal{R}(A)} = \mathcal{R}^{\perp}(A) \oplus \mathcal{N}^{\perp}(A^*) = \mathcal{R}^{\perp}(A) \oplus \overline{\mathcal{R}(A)}.\end{gathered} \tag{2.1}$$

Условимся далее обозначать принадлежность линейным подпространствам $\mathcal{N}(A)$ и $\mathcal{N}(A^*)$ с помощью верхнего индекса “0”, а принадлежность линейным подпространствам $\mathcal{N}^\perp(A)$ и $\mathcal{N}^\perp(A^*)$ — с помощью верхнего индекса “$\perp$”, таким образом, равенства $x = x^0 + x^\perp$ и $y = y^0 + y^\perp$ для элементов $x \in \mathcal{R}$ и $y \in \widetilde{\mathcal{R}}$ будут означать, что эти элементы разложены в соответствии с любым из представлений (2.1), но при этом всегда

$$x^0 \in \mathcal{N}(A) = \mathcal{R}^\perp(A^*), \qquad x^\perp \in \mathcal{N}^\perp(A) = \overline{\mathcal{R}(A^*)},$$
$$y^0 \in \mathcal{N}(A^*) = \mathcal{R}^\perp(A), \qquad y^\perp \in \mathcal{N}^\perp(A^*) = \overline{\mathcal{R}(A)}.$$

В ряде случаев ортогональные составляющие $x^\perp$ или $y^\perp$ могут принадлежать более узкому подмножеству, чем линейное подпространство $\mathcal{N}^\perp(A)$ или $\mathcal{N}^\perp(A^*)$ соответственно, и этот факт будет оговариваться особо.

Выделим в $\widetilde{\mathcal{R}}$ множество

$$\begin{aligned}\widetilde{\mathcal{D}} &= \mathcal{R}(A) \oplus \mathcal{N}(A^*) = \\ &= \big\{y = y^\perp + y^0\colon\ y^\perp \in \mathcal{R}(A),\, y^0 \in \mathcal{N}(A^*)\big\} = \\ &= \big\{y = Ax + y^0\colon\ x \in \mathcal{D}(A),\, y^0 \in \mathcal{N}(A^*)\big\}.\end{aligned}$$

Возьмём произвольный $y \in \widetilde{\mathcal{D}}$ и разложим его: $y = y^\perp + y^0$, где $y^\perp \in \mathcal{R}(A)$. Составляющая $y^\perp \in \mathcal{R}(A)$ может иметь много прообразов в $\mathcal{D}(A)$, т. е. равенство $y^\perp = Ax_1 = Ax_2$ возможно при $x_1 \neq x_2$. Однако легко понять, что в этом случае с необходимостью $x_1 - x_2 \in \mathcal{N}(A)$. Таким образом, если мы теперь применим разложение (2.1) к элементам $x_{1,2} \in \mathcal{D}(A)$ и напишем $x_1 = x_1^\perp + x_1^0$ и $x_2 = x_2^\perp + x_2^0$, то

$$x_1 - x_2 = (x_1^\perp - x_2^\perp) + (x_1^0 - x_2^0) \in \mathcal{N}(A) \quad\Longrightarrow\quad x_1^\perp - x_2^\perp = 0.$$

Другими словами, для любого $y^\perp \in \mathcal{R}(A)$ все его прообразы в $\mathcal{D}(A)$ имеют одну и ту же составляющую $x^\perp \in \mathcal{N}^\perp(A) \cap \mathcal{D}(A)$. При этом, разумеется, имеет место равенство $y^\perp = A(x^\perp + x^0) = Ax^\perp$.

Таким образом, для любого элемента $y \in \widetilde{\mathcal{D}}$ найдутся единственные составляющие $y^0 \in \mathcal{N}(A^*)$ и $y = y^\perp \in \mathcal{R}(A)$ такие, что $y = y^\perp + y^0$, а для каждого $y^\perp \in \mathcal{R}(A)$ существует единственный элемент $x^\perp \in \mathcal{N}^\perp(A) \cap \mathcal{D}(A)$ такой, что $y^\perp = Ax^\perp$. В результате соответствие $y \mapsto x^\perp$ задаёт оператор.

Определение. Оператор A^-, заданный на множестве

$$\widetilde{\mathcal{D}} = \mathcal{R}(A) \oplus \mathcal{N}(A^*) = \mathcal{R}(A) \oplus \mathcal{R}^\perp(A) \tag{2.2}$$

равенством $x^\perp = A^- y$, называется *псевдообратным* к оператору A.

Напишем ещё раз цепочку равенств, задающих действие псевдообратного оператора: для $y \in \mathcal{R}(A) \oplus \mathcal{N}(A^*)$

$$A^- y = A^-(y^\perp + y^0) = A^-(Ax + y^0) = A^-(A(x^\perp + x^0) + y^0) = x^\perp, \tag{2.3}$$

где $y^\perp \in \mathcal{R}(A)$ $y^0 \in \mathcal{N}(A^*)$, $x \in \mathcal{D}(A)$, $x^\perp \in \mathcal{N}^\perp(A) \cap \mathcal{D}(A)$ и $x^0 \in \mathcal{N}(A)$.

2.2. Простейшие свойства псевдообратного оператора. Приведем свойства, вытекающие непосредственно из определения псевдообратного оператора.

Свойство 1.

Оператор A^- линеен, плотно определён и замкнут.

Доказательство. Покажем линейность. Множество $\mathcal{D}(A^-) = \mathcal{R}(A) \oplus \mathcal{N}(A^*)$ есть линейное многообразие в $\widetilde{\mathcal{R}}$, поскольку $\mathcal{R}(A)$ и $\mathcal{N}(A^*)$ суть линейные многообразия. Предположим, что элементы $y_{1,2} = Ax^\perp_{1,2} + y^0_{1,2} \in \mathcal{D}(A^-)$. Тогда в силу линейности подпространств $\mathcal{R}(A)$, $\mathcal{N}^\perp(A)$, $\mathcal{N}(A^*)$ и линейности оператора A

$$\begin{aligned} A^-(\alpha_1 y_1 + \alpha_2 y_2) &= A^-\big((\alpha_1 A x_1^\perp + \alpha_2 A x_2^\perp) + (\alpha_1 y_1^0 + \alpha_2 y_2^0)\big) = \\ &= A^-\big(A(\alpha_1 x_1^\perp + \alpha_2 x_2^\perp) + (\alpha_1 y_1^0 + \alpha_2 y_2^0)\big), \end{aligned}$$

Разложение элемента $\alpha_1 y_1 + \alpha_2 y_2$ в сумму ортогональных слагаемых единственно, поэтому составляющие этого элемента, которые лежат в $\mathcal{R}(A)$ и $\mathcal{N}(A^*)$, суть соответственно $A(\alpha_1 x_1^\perp + \alpha_2 x_2^\perp)$ и $\alpha_1 y_1^0 + \alpha_2 y_2^0$. В силу единственности соответствия $\mathcal{R}(A) \mapsto \mathcal{N}^\perp(A) \cap \mathcal{D}(A)$

$$A^-(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 x_1^\perp + \alpha_2 x_2^\perp = \alpha_1 A^- y_1 + \alpha_2 A^- y_2,$$

и оператор A^- линейный.

Покажем, что оператор A^- плотно определён. Рассмотрим произвольный элемент $y \in \widetilde{\mathcal{R}}$ и разложим его в соответствии с представлением $\widetilde{\mathcal{R}} = \overline{\mathcal{R}(A)} \oplus \mathcal{N}(A^*)$ как $y = y^\perp + y^0$, $y^\perp \in \overline{\mathcal{R}(A)}$. Найдется последовательность $\{Ax_n\}_n$ такая, что $y^\perp = \lim Ax_n$. Положим $y_n = Ax_n + y^0$, тогда в силу ортогональности слагаемых мы имеем

$$\|y - y_n\|^2 = \|y^\perp - Ax_n\|^2 \to 0, \qquad y_n = Ax_n + y^0 \in \mathcal{D}(A^-),$$

т. е. любой элемент $y \in \widetilde{\mathcal{R}}$ сколь угодно точно приближается элементами из $\mathcal{D}(A^-)$, и $\mathcal{D}(A^-)$ плотно в пространстве $\widetilde{\mathcal{R}}$.

Покажем замкнутость оператора A^-: выведем, что условия $\{y_n\} \subset \mathcal{D}(A^-)$ и $\langle y_n, A^- y_n \rangle \to \langle y, x \rangle$ влекут $\langle y, x \rangle \in \Gamma(A^-)$, т. е. $y \in \mathcal{D}(A^-)$ и $x = A^- y$.

Итак, пусть $y_n \to y$ и $A^- y_n \to x$. Выполним разложения элементов $y_n \in \mathcal{D}(A^-)$, $y \in \widetilde{\mathcal{R}}$ и $x \in \mathcal{R}$ по ортогональным составляющим:

$$\begin{aligned}
y_n &= Ax_n^\perp + y_n^0, && x_n^\perp \in \mathcal{N}^\perp(A) \cap \mathcal{D}(A), && y_n^0 \in \mathcal{N}(A^*),\\
y &= y^\perp + y^0, && y^\perp \in \overline{\mathcal{R}(A)}, && y^0 \in \mathcal{N}(A^*)\\
x &= x^\perp + x^0, && x^\perp \in \mathcal{N}^\perp(A), && x^0 \in \mathcal{N}(A).
\end{aligned}$$

Тогда в силу сходимостей $y_n \to y$ и $A^- y_n = x_n^\perp \to x$ получаем

$$\|y_n - y\|^2 = \|Ax_n^\perp - y^\perp\|^2 + \|y_n^0 - y^0\|^2 \to 0,$$
$$\|A^- y_n - x\|^2 = \|x_n^\perp - x\|^2 = \|x_n^\perp - x^\perp\|^2 + \|x^0\|^2 \to 0,$$

следовательно, $x_n^\perp \to x^\perp$, $Ax_n^\perp \to y^\perp$ и $x^0 = 0$ (кроме того, разумеется, $y_n^0 \to y^0$, но это в данном случае не имеет значения). Первые две из указанных сходимостей могут быть записаны как $\langle x_n^\perp, Ax_n^\perp \rangle \to \langle x^\perp, y^\perp \rangle$. В силу замкнутости оператора A отсюда следует, что что $x^\perp \in \mathcal{D}(A)$ и $y^\perp = Ax^\perp$. Таким образом, предельный элемент $y = Ax^\perp + y^0 \in \mathcal{D}(A^-)$ и $x^\perp \in \mathcal{N}^\perp(A) \cap \mathcal{D}(A)$, поэтому $A^- y = x^\perp = x - x^0 = x$. Замкнутость псевдообратного оператора доказана.

Свойство 2.
Нуль-пространство и пространство значений оператора A^- суть

$$\mathcal{N}(A^-) = \mathcal{N}(A^*), \qquad \mathcal{R}(A^-) = \mathcal{N}^\perp(A) \cap \mathcal{D}(A). \tag{2.4}$$

Доказательство. Рассмотрим первое равенство. Если $y \in \mathcal{N}(A^*)$, то в силу единственности разложения $y = y^\perp + y^0$ следует написать $y = A0 + y$, где нулевой элемент, очевидно, принадлежит $\mathcal{N}^\perp(A) \cap \mathcal{D}(A)$, что означает $A^- y = 0$. Таким образом, $\mathcal{N}(A^*) \subset \mathcal{N}(A^-)$. С другой стороны, пусть $A^- y = 0$ для некоторого $y \in \mathcal{R}(A) \oplus \mathcal{N}(A^*)$, тогда $y = Ax^\perp + y^0$, причём $x^\perp = A^- y = 0$. Таким образом, $y = y^0 \in \mathcal{N}(A^*)$ и $\mathcal{N}(A^-) \subset \mathcal{N}(A^*)$. Первое равенство доказано.

Докажем второе равенство. Включение $\mathcal{R}(A^-) \subset \mathcal{N}^\perp(A) \cap \mathcal{D}(A)$ вытекает непосредственно из определения оператора A^-. Покажем обратное включение. Пусть элемент $x \in \mathcal{N}^\perp(A) \cap \mathcal{D}(A)$, тогда $Ax \in \mathcal{R}(A) \subset \mathcal{D}(A^-)$, и на Ax можно подействовать псевдообратным оператором. Но в этом случае $A^- Ax = A^-(Ax + 0) = x$, поскольку x принадлежит $\mathcal{N}^\perp(A)$. Таким образом, x есть результат действия оператора A^- на элемент Ax, т. е. $x \in \mathcal{R}(A^-)$.

Свойство 3.
Имеют место операторные равенства

$$AA^- A = A, \qquad A^- AA^- = A^-. \tag{2.5}$$

Доказательство. Прежде всего докажем, что $\mathcal{D}(AA^-A) = \mathcal{D}(A)$, т. е. для любого $x \in \mathcal{D}(A)$ определен элемент AA^-Ax. Имеем для $x \in \mathcal{D}(A)$ цепочку включений

$$
\begin{gathered}
Ax \in \mathcal{R}(A) \subset \mathcal{D}(A^-) = \mathcal{R}(A) \oplus \mathcal{N}(A^*), \\
A^-Ax \in \mathcal{R}(A^-) = \mathcal{N}^\perp(A) \cap \mathcal{D}(A) \subset \mathcal{D}(A).
\end{gathered}
$$

Далее для любого $x \in \mathcal{D}(A)$ проведем стандартное разложение $x = x^\perp + x^0$, где $x^\perp \in \mathcal{N}^\perp(A) \cap \mathcal{D}(A)$, тогда $Ax = Ax^\perp$. С другой стороны, $A^-Ax = A^-Ax^\perp = x^\perp$ по определению псевдообращения, следовательно, $AA^-Ax = Ax^\perp = Ax$. Первое равенство в (2.5) доказано. Перейдём к доказательству второго равенства.

Снова начнём с доказательства совпадения областей $\mathcal{D}(A^-AA^-) = \mathcal{D}(A)$. Пусть $y \in \mathcal{D}(A^-)$, тогда $A^-y \in \mathcal{N}^\perp(A) \cap \mathcal{D}(A) \subset \mathcal{D}(A)$, следовательно, определён элемент $AA^-y \in \mathcal{R}(A)$. Но $\mathcal{R}(A) \subset \mathcal{R}(A) \oplus \mathcal{N}(A^*) = \mathcal{D}(A^-)$, поэтому существует A^-AA^-y, тем самым $\mathcal{D}(A^-AA^-) = \mathcal{D}(A^-)$. Запишем разложение $y = Ax^\perp + y^0$, в котором $y^0 \in \mathcal{N}(A^*) = \mathcal{N}(A^-)$ (см. свойство 2). Тогда

$$
A^-y = A^-(Ax^\perp + y^0) = x^\perp, \qquad AA^-y = Ax^\perp, \qquad A^-AA^-y = A^-Ax^\perp = x^\perp
$$

по определению псевдообратного оператора. Итак, $A^-y = A^-AA^-y = x^\perp$ для любого $y \in \mathcal{D}(A^-)$, и мы получаем второе соотношение в (2.5).

Свойство 4.
Оператор $\overline{AA^-}$ есть ортогональный проектор на $\overline{\mathcal{R}(A)}$; оператор $\overline{A^-A}$ есть ортогональный проектор на $\overline{\mathcal{R}(A^*)}$.

Доказательство. Покажем, что оператор AA^- ограничен на $\mathcal{D}(A^-)$. Разложим произвольный элемент $y \in \mathcal{R}(A) \oplus \mathcal{N}(A^*)$ на соответствующие ортогональные составляющие: $y = y^\perp + y^0 = Ax + y^0$. В силу включения $y^0 \in \mathcal{N}(A^*) = \mathcal{N}(A^-)$ и равенств (2.5) мы имеем

$$
AA^-y = AA^-(Ax + y^0) = AA^-Ax = Ax,
$$

следовательно,

$$
\|y\|^2 = \|Ax\|^2 + \|y^0\|^2 = \|AA^-y\|^2 + \|y^0\|^2 \geqslant \|AA^-y\|^2.
$$

Таким образом, $\|AA^-y\| \leqslant \|y\|$ для любого $y \in \mathcal{D}(A^-)$. Отсюда вытекает ограниченность оператора AA^-. Замкнём данный оператор по непрерывности: для любого $y \in \overline{\mathcal{D}(A^-)} = \widetilde{\mathcal{R}}$ найдём последовательность $\{y_n\} \subset \mathcal{D}(A^-)$, сходящуюся к y, и положим $\overline{AA^-}y = \lim AA^-y_n$. При этом можно записать

$$
\begin{aligned}
&y = y^\perp + y^0, && y^\perp \in \overline{\mathcal{R}(A)}, && y^0 \in \mathcal{N}(A^*), \\
&y_n = Ax_n^\perp + y_n^0, && x_n^\perp \in \mathcal{N}^\perp(A) \cap \mathcal{D}(A), && y_n^0 \in \mathcal{N}(A^*),
\end{aligned}
$$

причем, очевидно, имеет место сходимость по каждой из ортогональных составляющих: $Ax_n^\perp \to y^\perp$ и $y_n^0 \to y^0$. С другой стороны, $A^- y_n = x_n^\perp$ и $AA^- y_n = Ax_n^\perp$, поэтому для любого $y \in \mathcal{R}$

$$\overline{AA^-} y \stackrel{\text{def}}{=} \lim_{n\to\infty} AA^- y_n = \lim_{n\to\infty} Ax_n^\perp = y^\perp,$$

где $y^\perp$ — составляющая элемента y в $\overline{\mathcal{R}(A)}$. Это как раз и означает, что $\overline{AA^-}$ есть ортогональный проектор на $\overline{\mathcal{R}(A)}$.

Аналогично, для любого $x \in \mathcal{D}(A)$, записав разложение $x = x^\perp + x^0$, имеем $A^- Ax = A^- Ax^\perp = x^\perp$, следовательно, $\|A^- Ax\|^2 = \|x^\perp\|^2 = \|x\|^2 - \|x^0\|^2 \leqslant \|x\|^2$, поэтому оператор $A^- A$ ограничен на $\mathcal{D}(A)$. Доопределим его по непрерывности (замкнём). Для любого $x \in \mathcal{R}$ пишем разложение $x = x^\perp + x^0$ и приближаем составляющую $x^\perp \in \mathcal{N}^\perp(A)$ элементами $x_n \in \mathcal{N}^\perp(A) \cap \mathcal{D}(A) = \mathcal{R}(A^-)$. Тогда $x_n = A^- y_n$ при некотором $y_n \in \mathcal{D}(A^-)$. В силу равенства $A^- AA^- = A^-$ получаем

$$\overline{A^- A} x \stackrel{\text{def}}{=} \lim_{n\to\infty} A^- A(x_n + x^0) = \lim_{n\to\infty} (A^- AA^- y_n + 0) = \lim_{n\to\infty} A^- y_n = x^\perp.$$

Таким образом, оператор $\overline{A^- A}$ есть ортогональный проектор на $\mathcal{N}^\perp(A) = \mathcal{R}(A^*)$. Свойство доказано.

Вычитая получившиеся проекторы из единичных операторов, действующих в соответствующих пространствах, получаем, что для любого $y \in \widetilde{\mathcal{R}}$ и любого $x \in \mathcal{R}$

$$(\tilde{I} - \overline{AA^-})y = y - y^\perp = y^0, \qquad (I - \overline{A^- A})x = x - x^\perp = x^0;$$

здесь и далее через I и $\tilde{I}$ мы обозначаем единичные операторы в пространствах $\mathcal{R}$ и $\widetilde{\mathcal{R}}$ соответственно. Таким образом, оператор $\overline{I - AA^-} = I - \overline{AA^-}$ есть ортогональный проектор на $\mathcal{N}(A^*) = \mathcal{R}^\perp(A)$, а $\overline{I - A^- A} = I - \overline{A^- A}$ — ортогональный проектор на $\mathcal{N}(A)$.

Вспомним, что если $A \in \mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, то существует сопряжённый оператор A^*, причём $A^* \in \mathbb{CL}(\widetilde{\mathcal{R}} \mapsto \mathcal{R})$. Таким образом, можно построить и оператор $(A^*)^-$, псевдообратный к A^*. Применяя свойство 4 к оператору A^* и его псевдообратному, получаем, что $\overline{(A^*)^- A^*}$ — ортогональный проектор на $\mathcal{R}(A^*)$, следовательно, $\overline{AA^-} = \overline{(A^*)^- A^*}$, потому что оба этих оператора проецируют на одно и то же подпространство. Аналогичные равенства можно записать и для всех остальных проекторов. Сведём все получившиеся результаты в таблицу.

Ортогональный проектор	Пространство, на которое он проецирует
$\overline{AA^-}$	$\overline{\mathcal{R}(A)}$
$\overline{A^-A}$	$\overline{\mathcal{R}(A^*)}$
$\overline{(A^*)^-A^*}$	$\overline{\mathcal{R}(A)}$
$\overline{A^*(A^*)^-}$	$\overline{\mathcal{R}(A^*)}$
$\tilde{I} - \overline{AA^-}$	$\mathcal{N}(A^*)$
$I - \overline{A^-A}$	$\mathcal{N}(A)$
$\tilde{I} - \overline{(A^*)^-A^*}$	$\mathcal{N}(A^*)$
$I - \overline{A^*(A^*)^-}$	$\mathcal{N}(A)$

Заметим, что если снять операцию замыкания с операторов в левом столбце таблицы, то AA^-, A^-A, $(A^*)^-A^*$ и $A^*(A^*)^-$ также будут ортогональными проекторами на своих областях определения $\mathcal{D}(A^-)$, $\mathcal{D}(A)$, $\mathcal{D}(A^*)$ и $\mathcal{D}(A^*)^-$ соответственно. Так, например, $AA^-y = y^\perp$ для любого $y \in \mathcal{D}(A^-) = \mathcal{R}(A) \oplus \mathcal{N}(A^*)$, где $y^\perp$ — составляющая элемента y в $\mathcal{R}(A)$; для любого $x \in \mathcal{D}(A)$ мы имеем $A^-Ax = x^\perp$, где $x^\perp$ — составляющая элемента x, лежащая в $\mathcal{N}^\perp(A) \cap \mathcal{D}(A)$. Перепишем нашу таблицу для операторов, определённых на незамкнутых линейных многообразиях.

Оператор	Область определения	Многообразие, на которое он проецирует
AA^-	$\mathcal{R}(A) \oplus \mathcal{N}(A^*)$	$\mathcal{R}(A)$
A^-A	$\mathcal{D}(A)$	$\mathcal{N}^\perp(A) \cap \mathcal{D}(A)$
$(A^*)^-A^*$	$\mathcal{D}(A^*)$	$\mathcal{N}^\perp(A^*) \cap \mathcal{D}(A^*)$
$A^*(A^*)^-$	$\mathcal{R}(A^*) \oplus \mathcal{N}(A)$	$\mathcal{R}(A^*)$
$\tilde{I} - AA^-$	$\mathcal{R}(A) \oplus \mathcal{N}(A^*)$	$\mathcal{R}^\perp(A)$
$I - A^-A$	$\mathcal{D}(A)$	$\mathcal{N}(A)$
$\tilde{I} - (A^*)^-A^*$	$\mathcal{D}(A^*)$	$\mathcal{N}(A^*)$
$I - A^*(A^*)^-$	$\mathcal{R}(A^*) \oplus \mathcal{N}(A)$	$\mathcal{R}^\perp(A^*)$

Теперь заметим, что $A^- \in \mathbb{CL}(\widetilde{\mathcal{R}} \mapsto \mathcal{R})$ по свойству 1, поэтому существует оператор $(A^-)^-$. Прежде чем доказывать, что он совпадает с A, покажем, что имеют место следующие равенства для линейных многообразий:

$$\mathcal{R}^\perp(A^-) = \mathcal{N}(A), \qquad \mathcal{R}(A^-)^- = \mathcal{R}(A). \tag{2.6}$$

Мы имеем в силу общих свойств ортогонального дополнения и свойства 2

$$\mathcal{R}^\perp(A^-) = \left(\overline{\mathcal{R}(A^-)}\right)^\perp = \left(\overline{\mathcal{N}^\perp(A) \cap \mathcal{D}(A)}\right)^\perp. \tag{2.7}$$

При этом
$$\overline{\mathcal{N}^{\perp}(A)\cap\mathcal{D}(A)}=\mathcal{N}^{\perp}(A). \tag{2.8}$$
В самом деле, если $x\in\overline{\mathcal{N}^{\perp}(A)\cap\mathcal{D}(A)}$, то $x=\lim x_n$, где $x_n\in\mathcal{N}^{\perp}(A)\cap\mathcal{D}(A)$, следовательно, $x_n\in\mathcal{N}^{\perp}(A)$ и $x=\lim x_n\in\overline{\mathcal{N}^{\perp}(A)}$. У замкнутого оператора пространство нулей замкнуто, поэтому мы показали, что из $x\in\overline{\mathcal{N}^{\perp}(A)\cap\mathcal{D}(A)}$ следует $x\in\mathcal{N}(A)$. Верно и обратное, если $x\in\mathcal{N}^{\perp}(A)$, то этот элемент, как и любой элемент из $\mathcal{R}$, может быть представлен как $x=\lim x_n$, где $x_n\in\mathcal{D}(A)$, поскольку $\mathcal{D}(A)$ плотно в $\mathcal{R}$. Напишем разложение $x_n=x_n^{\perp}+x_n^0$, где $x_n^{\perp}\in\mathcal{N}^{\perp}(A)\cap\mathcal{D}(A)$. Тогда в силу $x\in\mathcal{N}^{\perp}(A)$
$$\|x-x_n\|^2=\|x-x_n^{\perp}\|^2+\|x_n^0\|^2\to 0 \quad\Longrightarrow\quad \|x-x_n^{\perp}\|^2\to 0,\quad \|x_n^0\|^2\to 0,$$
другими словами, $x=\lim x_n^{\perp}$, $x_n^{\perp}\in\mathcal{N}^{\perp}(A)\cap\mathcal{D}(A)$, т. е. $x\in\overline{\mathcal{N}^{\perp}(A)\cap\mathcal{D}(A)}$. Равенство (2.8) доказано.

Заменяя в свойстве 2 (см. второе равенство в (2.4)) оператор A на A^-, получаем
$$\mathcal{R}(A^-)^-=\mathcal{N}^{\perp}(A^-)\cap\mathcal{D}(A^-)=\mathcal{N}^{\perp}(A^*)\cap\big(\mathcal{R}(A)\oplus\mathcal{N}(A^*)\big)=\overline{\mathcal{R}(A)}\cap\big(\mathcal{R}(A)\oplus\mathcal{N}(A^*)\big).$$
Если $y\in\mathcal{R}(A)$, то $y\in\overline{\mathcal{R}(A)}$ и $y\in\mathcal{R}(A)\oplus\mathcal{N}(A^*)$, следовательно
$$\mathcal{R}(A)\subset\overline{\mathcal{R}(A)}\cap\big(\mathcal{R}(A)\oplus\mathcal{N}(A^*)\big).$$
С другой стороны, если $y\notin\mathcal{R}(A)$, то $y=y^{\perp}+y^0$, $y^{\perp}\in\overline{\mathcal{R}(A)}$, $y^0\in\mathcal{N}(A^*)$, где либо $y^0\neq 0$, и тогда $y\notin\mathcal{N}^{\perp}(A^*)=\overline{\mathcal{R}(A)}$, либо $y^0=0$, но $y^{\perp}\notin\mathcal{R}(A)$, и тогда $y=y^{\perp}\notin\mathcal{R}(A)\oplus\mathcal{N}(A^*)$. Таким образом, $y\notin\overline{\mathcal{R}(A)}\cap\big(\mathcal{R}(A)\oplus\mathcal{N}(A^*)\big)$. Второе равенство в (2.6) доказано.

Свойство 5.

Дважды применённое псевдообращение возвращает нас к исходному оператору: $(A^-)^-=A$.

Доказательство. Заменяя в (2.2) оператор A на A^-, учитывая свойство 2 псевдообращения и первое равенство в (2.6), запишем
$$\mathcal{D}(A^-)^-=\mathcal{R}(A^-)\oplus\mathcal{R}^{\perp}(A^-)=\big(\mathcal{N}^{\perp}(A^-)\cap\mathcal{D}(A^-)\big)\oplus\mathcal{N}(A)=\mathcal{D}(A^-).$$
Пусть элемент $x\in\mathcal{D}(A^-)^-$, тогда
$$x=A^-y^{\perp}+x^0,\quad y^{\perp}=(A^-)^-x,\qquad x^0\in\mathcal{R}^{\perp}(A^-)=\mathcal{N}(A),\quad y^{\perp}\in\mathcal{R}(A^-)^-=\mathcal{R}(A),$$
где мы воспользовались равенствами (2.6). Поскольку элемент $y^{\perp}\in\mathcal{R}(A)$, найдется (единственный) элемент $x^{\perp}\in\mathcal{N}^{\perp}(A)\cap\mathcal{D}(A)$, при котором справедливо равенство $y^{\perp}=Ax^{\perp}$. Отсюда в соответствии со свойством $AA^-A=A$ имеем
$$Ax=A(A^-y^{\perp}+x^0)=A(A^-Ax^{\perp}+x^0)=AA^-Ax^{\perp}=Ax^{\perp}=y^{\perp}=(A^-)^-x.$$

Итак, $\mathcal{D}(A^-) = \mathcal{D}(A^{-1})$ и $A^-y = A^{-1}y$ для любого $y \in \mathcal{D}(A^-) = \mathcal{D}(A^{-1})$.

Свойство 6.

Если A — обратимый оператор, то $A^-y = A^{-1}y$ при $y \in \mathcal{D}(A^{-1}) = \mathcal{R}(A)$. Если к тому же A^* — обратимый оператор, то $A^- = A^{-1}$.

Доказательство. Понятно, что $\mathcal{D}(A^{-1}) = \mathcal{R}(A) \subset \mathcal{R}(A) \oplus \mathcal{N}(A^*) = \mathcal{D}(A^-)$. Обратимость A, т. е. условие $\mathcal{N}(A) = \{0\}$ гарантирует, что для любого $y \in \mathcal{R}(A)$ прообраз $x \in \mathcal{D}(A)$ такой, что $y = Ax$, единствен. При этом у элемента x составляющая в $\mathcal{N}(A) = \{0\}$ есть нулевой элемент, другими словами, $x \in \mathcal{N}^\perp(A) \cap \mathcal{D}(A)$. Отсюда $A^-y = x = A^{-1}y$.

В общем случае область $\mathcal{D}(A^-) = \mathcal{R}(A) \oplus \mathcal{N}(A^*)$ шире, чем $\mathcal{D}(A^{-1}) = \mathcal{R}(A)$, однако если A^* — обратимый оператор, то $\mathcal{R}^\perp(A) = \mathcal{N}(A^*) = \{0\}$. В этом случае $\mathcal{D}(A^-) = \mathcal{R}(A) = \mathcal{D}(A^{-1}) = \mathcal{D}(A^-)$, и обратный оператор совпадает с псевдообратным.

Замечание 1. Последнее утверждение показывает, что $A^- = A^{-1}$ в том и только том случае, когда обратимы A и A^*.

Поскольку операторы A^- и A^* замкнуты и плотно определены, существуют операторы $(A^-)^*$ и $(A^*)^-$. Эти операторы совпадают, но полное доказательство этого факта требует привлечения дополнительного материала, который мы рассмотрим ниже. Сейчас мы докажем не равенство $(A^-)^* = (A^*)^-$, а включение графиков $\Gamma(A^-)^* \subset \Gamma(A^*)^-$.

Свойство 7.

Имеют место включение $\mathcal{D}(A^*)^- \subset \mathcal{D}(A^-)^*$ и равенство $(A^*)^-x = (A^-)^*x$ для любого $x \in \mathcal{D}(A^*)^-$.

Доказательство. По определению

$$\begin{aligned} &\mathcal{D}(A^*)^- = \mathcal{R}(A^*) \oplus \mathcal{R}^\perp(A^*), \\ &\mathcal{D}(A^-)^* = \Big\{x \in \mathcal{R}\colon \sup_{\substack{u \in \mathcal{D}(A^-),\\ u \neq 0}} \frac{|(A^-u, x)|}{\|u\|} < \infty\Big\}. \end{aligned} \tag{2.9}$$

Пусть $x \in \mathcal{D}(A^*)^- = \mathcal{R}(A^*) \oplus \mathcal{R}^\perp(A^*)$ и $u \in \mathcal{D}(A^-) = \mathcal{R}(A) \oplus \mathcal{N}(A^*)$ — произвольные элементы. Разложим их на ортогональные составляющие:

$$\begin{aligned} &x = A^*y^\perp + x^0, \qquad && y^\perp \in \mathcal{N}^\perp(A^*) \cap \mathcal{D}^\perp(A^*), \qquad && x^0 \in \mathcal{N}(A) = \mathcal{R}^\perp(A^*); \\ &u = Az^\perp + u^0, && z^\perp \in \mathcal{N}^\perp(A) \cap \mathcal{D}(A), && u^0 \in \mathcal{N}(A^*) = \mathcal{R}^\perp(A). \end{aligned}$$

Тогда $A^-u = z^\perp$, кроме того, в силу ортогональности составляющих с учётом равенства $A^{**} = A$ и включения $z^\perp \in \mathcal{D}(A)$

$$(A^-u, x) = (z^\perp, A^*y^\perp + x^0) = (z^\perp, A^*y^\perp) = (Az^\perp, y^\perp) = \\ = (Az^\perp + u^0, y^\perp) = (u, y^\perp).$$

Отсюда следует, что $|(A^-u, x)| \leqslant \|u\| \cdot \|y^\perp\|$ и для любого $x = A^*y^\perp + x^0 \in \mathcal{D}(A^*)^-$

$$\sup_{\substack{u\in\mathcal{D}(A^-),\\ u\neq 0}} \frac{|(A^-u, x)|}{\|u\|} \leqslant \|y^\perp\| < \infty,$$

т. е. $x \in \mathcal{D}(A^-)^*$, что означает $\mathcal{D}(A^*)^- \subset \mathcal{D}(A^-)^*$.

Далее, $(A^*)^-x = y^\perp$, поэтому для $u = Az^\perp + u^0 \in \mathcal{D}(A^-)$

$$(u, (A^*)^-x) = (u, y^\perp) = (Az^\perp + u^0, y^\perp) = (Az^\perp, y^\perp),$$
$$(u, (A^-)^*x) = (A^-u, x) = (z^\perp, A^*y^\perp + x^0) = (z^\perp, A^*y^\perp) = (Az^\perp, y^\perp),$$

где мы учли, что $A^- \in \mathbb{CL}(\widetilde{\mathcal{R}} \mapsto \mathcal{R})$, поэтому $(A^-)^{**} = A^-$. Следовательно, мы можем утверждать, что $(u, (A^-)^*x) = (u, (A^*)^-x)$ для любого $u \in \mathcal{D}(A^-)$. Множество $\mathcal{D}(A^-)$ плотно в $\widetilde{\mathcal{R}}$, поэтому последнее равенство влечёт $(A^*)^-x = (A^-)^*x$. Таким образом, $\Gamma(A^-)^* \subset \Gamma(A^*)^-$. Обратное включение и тем самым равенство $(A^*)^- = (A^-)^*$ будет показано далее.

2.3. Псевдообращение и вариационные задачи. В этом разделе мы покажем, как используется псевдобратный оператор для решения вариационных задач. Начнём с задачи, которая в линейной алгебре носит название «метод наименьших квадратов».

Задача 1. Пусть $A \in \mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ и элемент $y \in \widetilde{\mathcal{R}}$ произволен, но фиксирован. Требуется найти минимум по $x \in \mathcal{D}(A)$ функционала $\|Ax - y\|^2$ и множество элементов в $\mathcal{D}(A)$, на которых этот минимум достигается.

Теорема 3.
Пусть $A \in \mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ и $y \in \widetilde{\mathcal{R}}$. Тогда

$$\inf_{x\in\mathcal{D}(A)} \|Ax - y\|^2 = \|y^0\|^2 = \|(I - \overline{AA^-})y\|^2. \tag{2.10}$$

Если $y \in \mathcal{D}(A^-)$, то множество элементов, на которых достигается точная нижняя грань, имеет вид

$$\mathcal{D}_*(y) = \big\{x = A^-y + x^0,\ x^0 \in \mathcal{N}(A)\big\}. \tag{2.11}$$

Если $y \notin \mathcal{D}(A^-)$, то множество $\mathcal{D}_*(y)$ пустое и всякая минимизирующая последовательность вариационной задачи (2.10) неограничена.

Доказательство. Разложим произвольный элемент $y \in \widetilde{\mathcal{R}}$ на составляющие стандартным образом: $y = y^\perp + y^0$, $y^\perp \in \overline{\mathcal{R}(A)}$, $y^0 \in \mathcal{R}^\perp(A)$. Тогда

$$\|Ax - y\|^2 = \|Ax - y^\perp\|^2 + \|y^0\|^2 \geqslant \|y^0\|^2 = \|(I - \overline{AA^-})y\|^2,$$

где в последнем равенстве мы воспользовались свойством 4 псевдообратного оператора.

Если и только если $y^\perp \in \mathcal{R}(A)$ найдётся $x \in \mathcal{D}(A)$ такой, что $Ax = y^\perp$. Если же $y^\perp \in \overline{\mathcal{R}(A)} \backslash \mathcal{R}(A)$, то равенство $\|Ax - y^\perp\|^2 = 0$ невозможно, но найдётся последовательность $\{y_n\} \subset \mathcal{R}(A)$ (т. е. $y_n = Ax_n$), для которой $y^\perp = \lim y_n$. Другими словами, найдётся последовательность $\{x_n\} \subset \mathcal{D}(A)$ такая, что $\|y^\perp - Ax_n\| \to 0$. Таким образом, мы показали, что имеет место соотношение (2.10), а также что множество $\mathcal{D}_*(y)$ непусто тогда и только тогда, когда $y^\perp \in \mathcal{R}(A)$, т. е. тогда и только тогда, когда $y = y^\perp + y^0 \in \mathcal{D}(A^-) = \mathcal{R}(A) \oplus \mathcal{R}^\perp(A)$.

Для доказательства (2.11) заметим, что если $y = y^\perp + y^0 = Ax^\perp + y^0 \in \mathcal{D}(A^-)$, то элемент $x^\perp = A^- y$ даёт $Ax^\perp = y^\perp$ и $\|Ax^\perp - y\|^2 = \|y^0\|^2$, т. е. $x^\perp \in \mathcal{D}_*(y)$. Если элемент x имеет вид $x = x^\perp + x^0$, где $x^0 \in \mathcal{N}(A)$, то, очевидно, $Ax^\perp = Ax = y^\perp$, следовательно, $x \in \mathcal{D}_*(y)$. Верно и обратное, если $x \in \mathcal{D}_*(y)$, то $\|Ax - y\|^2 = \|y^0\|^2$, следовательно, $Ax = y^\perp$, другими словами, $Ax = Ax^\perp$. Отсюда $x - x^\perp \in \mathcal{N}(A)$, т. е. $x = x^\perp + x^0$ при некотором $x^0 \in \mathcal{N}(A)$. Равенство (2.11) доказано.

Пусть теперь $y \notin \mathcal{D}(A^-)$, т. е. $y^\perp \in \overline{\mathcal{R}(A)} \setminus \mathcal{R}(A)$. Докажем, что любая последовательность $\{x_n\} \subset \mathcal{D}(A)$ такая, что $\|Ax_n - y\|^2 \to \|y^0\|^2$, с необходимостью неограничена.

Предположим обратное: допустим, что минимизирующая последовательность $\{x_n\}$ ограничена. Поскольку имеет место сходимость $\|Ax_n - y\|^2 \to \|y^0\|^2$, последовательность $\{\|Ax_n - y\|\}_{n=\overline{1,\infty}}$ ограничена, следовательно (ибо $\|Ax_n\| \leqslant \|Ax_n - y\| + \|y\|$), также ограничена последовательность $\{Ax_n\}$. Итак, ограничены обе последовательности, $\{x_n\}$ и $\{Ax_n\}$. Согласно теореме о слабой компактности (слабо) ограниченной последовательности (см. приложение B, теорема B.4) из $\{x_n\}$ можно выделить слабо сходящуюся подпоследовательность $\{x_{n_k}\}_{k=\overline{1,\infty}}$. Далее из ограниченной последовательности $\{Ax_{n_k}\}_{k=\overline{1,\infty}}$ тоже выделяется слабо сходящаяся подпоследовательность $\{Ax_{n_{k_j}}\}_{j=\overline{1,\infty}}$. Для удобства обозначений положим $x_{n_{k_j}} = z_j$. Тогда можно записать $z_j \xrightarrow{\text{w}} z$ и $Az_j \xrightarrow{\text{w}} u$, имея в виду слабую сходимость. Это означает, что $\langle z_j, Az_j\rangle \xrightarrow{\text{w}} \langle z, u\rangle$. Из слабой замкнутости графика (см. приложение B, терема B.6) следует, что $\langle z, u\rangle \in \Gamma(A)$, другими словами, $u = Az$, $z \in \mathcal{D}(A)$.

Итак, с одной стороны, $\|Az_j - y\|^2 \to \|y^0\|^2$ (напомним, что $\{z_j\}$ — подпоследовательность минимизирующей последовательности $\{x_n\}$), с другой стороны, имеют место слабые сходимости $Az_j \xrightarrow{\text{w}} Az$ и, следовательно, $Az_j - y \xrightarrow{\text{w}} Az - y$. По теореме о слабой полунепрерывности нормы снизу (см. приложение B, теорема B.3)

$$\|Az - y\| \leqslant \varliminf_{j\to\infty} \|Az_j - y\| = \lim_{j\to\infty} \|Az_j - y\| = \|y^0\|.$$

Таким образом,

$$\|Az - y\| \leqslant \|y^0\| = \inf_{x\in\mathcal{D}(A)} \|Ax - y\|,$$

следовательно, элемент z — точка минимума в (2.10). Это противоречит тому, что минимум не достигается (напомним, $y \notin \mathcal{D}(A^-)$). Терема доказана.

Замечание 2. Если $y \in \mathcal{D}(A^-)$, то среди элементов, на которых достигается минимум в (2.10), элемент A^-y имеет минимальную норму, ибо для любого элемента $x \in \mathcal{D}_*(y)$ в силу ортогональности $x^\perp = A^-y$ и x^0 можно записать оценку $\|x\|^2 = \|A^-y\|^2 + \|x^0\|^2 \geqslant \|A^-y\|^2$. Этот факт может быть использован как альтернативное определение псевдообратного оператора. В этом случае за область определения $\mathcal{D}(A^-)$ принимают множество тех y, при которых $\mathcal{D}_*(y) \neq \varnothing$, и в качестве $x = A^-y$ берут элемент $x \in \mathcal{D}_*(y)$ с минимальной нормой. Легко показать, что множество $\mathcal{D}_*(y)$ выпукло и замкнуто, поэтому в нём существует только один элемент x_y с минимальной нормой. Таким образом, соответствие $y \mapsto x_y$ задаёт оператор.

Видно, что в случае $y \notin \mathcal{D}(A^-)$ попытка минимизировать невязку $\|Ax - y\|^2$ приводит к тому, что минимизирующая последовательность становится неограниченно большой (по норме). В этом случае естественной является идея минимизации невязки при одновременной «стабилизации» минимизирующей последовательности. Мы приходим к следующей задаче.

Задача 2. Пусть заданы оператор $A \in \mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, элемент $y \in \widetilde{\mathcal{R}}$ и параметр $\omega > 0$. Требуется найти множество элементов в $\mathcal{D}(A)$, на которых достигается минимум по $x \in \mathcal{D}(A)$ функционала $\|Ax - y\|^2 + \omega\|x\|^2$.

Замечание 3. Функционал $L(x) = \|Ax - y\|^2 + \omega\|x\|^2$ можно отождествить с функцией Лагранжа задачи на условный экстремум

$$\min\left\{\|Ax - y\|^2 \,\middle|\, x \in \mathcal{D}(A)\colon \|x\|^2 \leqslant C\right\},$$

где ω играет роль параметра Лагранжа

Теорема 4.

Для любых $\omega > 0$ и $y \in \widetilde{\mathcal{R}}$

$$\min_{x\in\mathcal{D}(A)} \left\{\|Ax-y\|^2+\omega\|x\|^2\right\} \tag{2.12}$$

достигается на единственном элементе $x_\omega = A^*(AA^*+\omega\tilde{I}\,)^{-1}y$.

Доказательство. Положим $L(x) = \|Ax-y\|^2+\omega\|x\|^2$, тогда тогда для любого $h \in \mathcal{D}(A)$

$$L(x+h)-L(x) = 2(Ah, Ax-y)+2\omega(x,h)+\|Ah\|^2+\omega\|h\|^2. \tag{2.13}$$

На основании теоремы фон Неймана (см. приложение В, замечание после теоремы В.8) можно утверждать, что операторы

$$(AA^*+\omega\tilde{I}\,)^{-1}, \qquad A^*(AA^*+\omega\tilde{I}\,)^{-1}, \qquad AA^*(AA^*+\omega\tilde{I}\,)^{-1}$$

всюду определены, поэтому элемент $x_\omega = A^*(AA^*+\omega\tilde{I}\,)^{-1}y$ существует и принадлежит $\mathcal{D}(A)$. Кроме того, $(AA^*+\omega\tilde{I})(AA^*+\omega\tilde{I}\,)^{-1} = \tilde{I}$, поэтому

$$Ax_\omega - y = AA^*(AA^*+\omega\tilde{I}\,)^{-1}y-(AA^*+\omega\tilde{I})(AA^*+\omega\tilde{I}\,)^{-1}y = -\omega(AA^*+\omega\tilde{I}\,)^{-1}y.$$

Отсюда $Ax_\omega - y \in \mathcal{D}(A^*)$ и

$$\begin{aligned}(Ah, Ax_\omega-y)+\omega(x_\omega,h) &= -\omega(Ah,(AA^*+\omega\tilde{I}\,)^{-1}y)+\omega(x_\omega,h) = \\ &= -\omega(h, A^*(AA^*+\omega\tilde{I}\,)^{-1}y)+\omega(A^*(AA^*+\omega\tilde{I}\,)^{-1}y,h) = 0.\end{aligned}$$

Подставляя это в (2.13), получаем, что для всех $h \in \mathcal{D}(A)$

$$L(x_\omega+h)-L(x_\omega) = \|Ah\|^2+\omega\|h\|^2 \geqslant 0,$$

и равенство $L(x_\omega+h)-L(x_\omega) = 0$ возможно, если и только если $h = 0$, откуда вытекает, что x_ω – единственная точка минимума функционала $L(\,\cdot\,)$. Теорема доказана.

Покажем, как связаны две рассмотренные нами задачи на минимум.

Теорема 5.

Пусть $y \in \widetilde{\mathcal{R}}$ произволен, $\omega_n \to +0$ и $x_n = A^*(AA^*+\omega_n\tilde{I}\,)^{-1}y$. Тогда

$$\inf_{x\in\mathcal{D}(A)} \|Ax-y\|^2 = \lim_{n\to\infty}\|Ax_n-y\|^2. \tag{2.14}$$

Доказательство. Напомним, что значение точной нижней грани в левой части равенства (2.14) равно $\|y^0\|^2$. Пусть $\varepsilon > 0$ произвольно. Найдём $x_\varepsilon \in \mathcal{D}(A)$, для которого справедливо неравенство $\|Ax_\varepsilon - y\|^2 \leqslant \|y^0\|^2 + \varepsilon/2$. Далее, вспомнив, что $\omega_n \to 0$, найдём номер $N = N(x_\varepsilon) = N(\varepsilon)$ такой, что $\omega_n \|x_\varepsilon\|^2 < \varepsilon/2$ при всех $n > N(\varepsilon)$. С другой стороны, имеем

$$\|Ax_n - y\|^2 \leqslant \|Ax_n - y\|^2 + \omega_n \|x_n\|^2 \leqslant \|Ax_\varepsilon - y\|^2 + \omega_n \|x_\varepsilon\|^2,$$

где второе неравенство вытекает из того, что x_n — точка минимума функционала $\|Ax - y\|^2 + \omega_n \|x\|^2$. Подставляя оценки для $\|Ax_\varepsilon - y\|^2$ и $\omega_n \|x_\varepsilon\|^2$, получаем

$$\|Ax_\varepsilon - y\|^2 + \omega_n \|x_\varepsilon\|^2 \leqslant \|y^0\|^2 + \varepsilon/2 + \varepsilon/2.$$

Таким образом, $\|Ax_n - y\|^2 \leqslant \|y^0\|^2 + \varepsilon$, если $n > N(\varepsilon)$, т. е. $\|Ax_n - y\|^2 \to \|y^0\|^2$. Теорема доказана.

Воспользуемся доказанными теоремами и завершим доказательство равенства $(A^*)^- = (A^-)^*$. Напомним, мы доказали включение $\mathcal{D}(A^*)^- \subset \mathcal{D}(A^-)^*$, т. е. показали, что если $x \in \mathcal{D}(A^*)^-$, то $x \in \mathcal{D}(A^-)^*$, и теперь нам требуется доказать обратное включение. Будем доказывать его от противного.

Пусть $x \notin \mathcal{D}(A^*)^-$. Рассмотрим $\|A^*y - x\|^2$ как функционал от $y \in \mathcal{D}(A^*)$. В силу теоремы 3 (с заменой A на A^*) можно утверждать, что $\inf_{y \in \mathcal{D}(A^*)} \|A^*y - x\|^2$ не достигается и любая минимизирующая последовательность неограничена. В частности, неограничена последовательность $y_n = A(A^*A + \omega_n I)^{-1}x$, $n = 1, 2 \ldots$, которая является минимизирующей в силу теоремы 5 (опять же мы должны заменить в этой теореме A на A^*). Другими словами, мы имеем

$$\|y_n\| = \left\|A(A^*A + \omega_n I)^{-1}x\right\| \underset{n\to\infty}{\longrightarrow} \infty. \tag{2.15}$$

Положим $z_n = \bigl(A^*A + \omega_n I\bigr)^{-1}x$, тогда $y_n = Az_n$ и $x = (A^*A + \omega_n I\bigr)z_n$, т. е.

$$x = A^*Az_n + \omega_n z_n. \tag{2.16}$$

Выполним разложение

$$z_n = z_n^\perp + z_n^0, \qquad z_n^\perp \in \mathcal{N}^\perp(A) \cap \mathcal{D}(A), \quad z_n^0 \in \mathcal{N}(A),$$

и с учётом того, что $x \notin \mathcal{D}(A^*)^- = \mathcal{R}(A^*) \oplus \mathcal{N}(A)$, — разложение

$$x = x^\perp + x^0, \qquad x^\perp \in \overline{\mathcal{R}(A^*)} \setminus \mathcal{R}(A^*), \qquad x^0 \in \mathcal{N}(A).$$

Подставим эти разложения в (2.16) и сгруппируем слагаемые, лежащие в одном подпространстве: поскольку $Az^0 = 0$, имеем

$$x^\perp + x^0 = (A^*Az_n^\perp + \omega_n z_n^\perp) + \omega_n z_n^0.$$

Отсюда

$$x^{\perp} = A^*Az_n^{\perp} + \omega_n z_n^{\perp}, \qquad x^0 = \omega_n z_n^0. \tag{2.17}$$

С другой стороны, $A^-y_n = A^-Az_n = A^-A(z_n^{\perp} + z_n^0) = A^-Az_n^{\perp} = z_n^{\perp}$, где последнее равенство следует из того, что оператор A^-A проецирует элемент $z_n \in \mathcal{D}(A)$ на многообразие $\overline{\mathcal{R}(A^*)} \cap \mathcal{D}(A)$. Следовательно, принимая во внимание первое равенство в (2.17) и очевидное равенство $Az_n^0 = 0$, имеем

$$\begin{aligned}(A^-y_n, x) = (z_n^{\perp}, x^{\perp} + x^0) = (z_n^{\perp}, x^{\perp}) = (z_n^{\perp}, A^*Az_n^{\perp} + \omega_n z_n^{\perp}) = \\ = \|Az_n^{\perp}\|^2 + \omega_n\|z_n^{\perp}\|^2 = \|A(z_n^{\perp} + z_n^0)\|^2 + \omega_n\|z_n^{\perp}\|^2 = \|Az_n\|^2 + \omega_n\|z_n^{\perp}\|^2 \geqslant \\ \geqslant \|Az_n\|^2 = \|y_n\|^2,\end{aligned}$$

При этом имеет место условие неограниченности (2.15). Таким образом,

$$\frac{(A^-y_n, x)}{\|y_n\|} \geqslant \|y_n\| \underset{n\to\infty}{\longrightarrow} \infty, \qquad \sup_{\substack{y\in\mathcal{D}(A^-),\\ y\neq 0}} \frac{|(A^-y, x)|}{\|y\|} = \infty,$$

следовательно, $x \notin \mathcal{D}(A^-)^*$.

Итак, мы доказали, что $x \notin \mathcal{D}(A^*)^-$ влечёт $x \notin \mathcal{D}(A^-)^*$, а это означает, что $\mathcal{D}(A^*)^- \subset \mathcal{D}(A^-)^*$; в совокупности с уже доказанным ранее обратным включением мы имеем $\mathcal{D}(A^*)^- = \mathcal{D}(A^-)^*$, и тем самым доказательство равенства $(A^*)^- = (A^-)^*$ завершено.

Доказанные теоремы позволяют получить несколько новых свойств псевдообратного оператора.

Свойство 8.

Для любого $y \in \widetilde{\mathcal{R}}$ и любой последовательности $\omega_n \to +0$

$$\lim_{n\to\infty} AA^*(AA^* + \omega_n I)^{-1}y = \overline{AA^-}\, y = \overline{A^*(A^*)^-}\, y. \tag{2.18}$$

Доказательство. Разложим произвольный элемент $y \in \widetilde{\mathcal{R}}$ на ортогональные составляющие: $y = y^{\perp} + y^0$, где

$$y^{\perp} = \overline{AA^-}y \in \overline{\mathcal{R}(A)}, \qquad y^0 = \overline{I - AA^-}y \in \mathcal{R}^{\perp}(A).$$

Тогда при $x_n = A^*(AA^* + \omega_n I)^{-1}y$ в силу теорем 3 и 5

$$\|y^0\|^2 = \inf_{x\in\mathcal{D}(A)} \|Ax - y\| = \lim_{n\to\infty} \|Ax_n - y\|^2 = \lim_{n\to\infty} \|Ax_n - y^{\perp}\|^2 + \|y^0\|^2,$$

откуда $\|Ax_n - y^{\perp}\| \to 0$, что эквивалентно $Ax_n \to y^{\perp}$ или с учётом введённых обозначений $AA^*(AA^* + \omega_n I)^{-1}y \to \overline{AA^-}y$. Первое равенство в (2.18) доказано.

Второе равенство в (2.18) вытекает из того, что $\overline{AA^-} = \overline{(A^*)^- A^*}$, поскольку и тот, и другой оператор проецируют на $\overline{\mathcal{R}(A)}$.

Свойство 9.
Для любого $y \in \mathcal{D}(A^-)$ и любой последовательности $\omega_n \to +0$

$$A^- y = \lim_{n\to\infty} A^*(AA^* + \omega_n I)^{-1} y. \tag{2.19}$$

Доказательство. Положим $y = Ax^\perp + y^0$, $x^\perp = A^- y$ и $x_n = A^*(AA^* + \omega_n I)^{-1} y$. Тогда по теореме 5 в силу условия $y \in \mathcal{D}(A^-)$

$$\|Ax_n - y\|^2 \to \|y^0\|^2 = \min_{x\in\mathcal{D}(A)} \|Ax - y\|^2. \tag{2.20}$$

Отсюда следует ограниченность последовательности $\{Ax_n - y\}$ и, следовательно, последовательности $\{Ax_n\}$. Далее, мы можем записать следующую цепочку неравенств:

$$\|Ax_n - y\|^2 + \omega_n\|x_n\|^2 \leqslant \|Ax^\perp - y\|^2 + \omega_n\|x^\perp\|^2 \leqslant \|Ax_n - y\|^2 + \omega_n\|x^\perp\|^2, \tag{2.21}$$

где в первом неравенства мы учли, что x_n является точкой минимума функционала $\|Ax - y\|^2 + \omega_n\|x\|^2$, а второе неравенство следует из того, что $x^\perp$ — точка минимума функционала $\|Ax-y\|^2$. Сравнивая левую и правую части (2.21), видим, что

$$\|x_n\|^2 \leqslant \|x^\perp\|^2, \tag{2.22}$$

и последовательность $\{x_n\}$ также ограничена. Теперь мы готовы показать, что имеет место сходимость $x_n \to x^\perp$.

Пусть $\{x_{n_k}\}$ – произвольная подпоследовательность в $\{x_n\}$. Эта подпоследовательность, а также соответствующая ей подпоследовательность $\{Ax_{n_k}\}$ ограничены. Тогда из них можно выделить слабо сходящиеся подпоследовательности. Чтобы не писать «многоэтажные» индексы, обозначим члены этой слабо сходящейся подпоследовательности через z_j, $j = 1, 2, \ldots$.

Итак, пусть для некоторых элементов $z \in \mathcal{R}$ и $u \in \widetilde{\mathcal{R}}$

$$z_j \xrightarrow{\mathrm{w}} z, \qquad Az_j \xrightarrow{\mathrm{w}} u. \tag{2.23}$$

Другими словами, $\langle z_j, Az_j\rangle \xrightarrow{\mathrm{w}} \langle z, u\rangle$, и в силу слабой замкнутости $\Gamma(A)$ мы заключаем, что $z \in \mathcal{D}(A)$ и $u = Az$ (мы рассуждали аналогично при доказательстве неограниченности для $y \notin \mathcal{D}(A^-)$ минимизирующей последовательности метода наименьших квадратов). Итак, $Az_j \xrightarrow{\mathrm{w}} Az$, следовательно, $Az_j - y \xrightarrow{\mathrm{w}} Az - y$.

Покажем, что с необходимостью $z = x^\perp$. Имеем для $\{z_j\} \subset \{x_n\}$ по теореме о слабой полунепрерывности нормы

$$\|Az - y\|^2 \leqslant \varliminf_{j\to\infty} \|Az_j - y\|^2 = \lim_{n\to\infty} \|Ax_n - y\|^2 = \min_{x\in\mathcal{D}(A)} \|Ax - y\|^2,$$

где мы учли (2.20). Таким образом, z — точка минимума функционала $\|Ax_n - y\|^2$, следовательно, $z = x^\perp + z^0$, где $z^0 \in \mathcal{N}(A)$ (см. формулу (2.11) для множества точек минимума). Но при этом с учётом (2.22) и (2.23)

$$\|x^\perp\|^2 + \|z^0\|^2 = \|z\|^2 \leqslant \varliminf_{j\to\infty} \|z_j\|^2 \leqslant \|x^\perp\|^2,$$

откуда $z^0 = 0$ и $z = x^\perp$.

Теперь легко показать, что данном случае из слабой сходимости вытекает сходимость последовательности $\{z_j\}$ к $x^\perp$ по норме: ещё раз запишем $\|z_j\| \leqslant \|x^\perp\|$, в результате имеем

$$\|z_j - x^\perp\|^2 = \|z_j\|^2 - 2(z_j, x^\perp) + \|x^\perp\|^2 \leqslant 2\|x^\perp\|^2 - 2(z_j, x^\perp) \to 2\|x^\perp\|^2 - 2(x^\perp, x^\perp) = 0,$$

т. е. $\|z_j - x^\perp\| \to 0$. Таким образом, мы показали, что в любой последовательности $\{x_{n_k}\} \subset \{x_n\}$ можно найти подпоследовательность, которая сходится к $x^\perp$.

Отсюда немедленно следует (2.19). В самом деле, если последовательность $\{x_n\}$ (т. е. последовательность $\{A^*(AA^* + \omega_n I)^{-1} y\}_{n=\overline{1,\infty}}$ согласно принятым обозначениям) не сходится к $x^\perp = A^- y$, то найдется $\varepsilon_0 > 0$ и подпоследовательность $\{x_{n_k}\}$ такие, что $\|x_{n_k} - x^\perp\| \geqslant \varepsilon_0$ для всех $k = 1, 2, \ldots$. Из такой подпоследовательности, разумеется, нельзя выделить подпоследовательность, сходящуюся к $x^\perp$. Полученное противоречие доказывает наше утверждение.

Заметим, что в силу теоремы фон Неймана оператор $(AA^* + \omega\tilde{I}\,)^{-1}$ всюду определен, следовательно, $(AA^* + \omega\tilde{I})(AA^* + \omega\tilde{I}\,)^{-1} = \tilde{I}$. Отсюда для любого $y \in \mathcal{D}(A^*)$

$$A^*y = A^*\big(AA^* + \omega\tilde{I}\,\big)\big(AA^* + \omega\tilde{I}\,\big)^{-1}y = \big(A^*A + \omega\tilde{I}\,\big)A^*\big(AA^* + \omega\tilde{I}\,\big)^{-1}y.$$

Подействуем на левую и правую части этого равенства оператором $\big(AA^* + \omega\tilde{I}\,\big)^{-1}$, получим

$$\big(A^*A + \omega I\,\big)^{-1}A^*y = A^*\big(AA^* + \omega I\,\big)^{-1}y. \tag{2.24}$$

По той же теореме фон Неймана оператор $\|A^*(AA^* + \omega I)^{-1}\| \leqslant \omega^{-1}$, следовательно,

$$\sup_{\substack{y\in\mathcal{D}(A^*),\\ \|y\|\leqslant 1}} \|\big(A^*A + \omega I\,\big)^{-1}A^*y\| = \sup_{\substack{y\in\mathcal{D}(A^*),\\ \|y\|\leqslant 1}} \|A^*\big(A^*A + \omega I\,\big)^{-1}y\| \leqslant \frac{1}{\omega},$$

т. е. оператор $\big(A^*A+\omega\tilde{I}\,\big)^{-1}A^*$ ограничен на $\mathcal{D}(A^*)$. Доопределяем его по непрерывности (замыкаем) на $\overline{\mathcal{D}(A^*)} = \widetilde{\mathcal{R}}$. Очевидно, в силу (2.24) имеет место следующее операторное равенство:

$$\overline{\big(A^*A+\omega\tilde{I}\,\big)^{-1}A^*} = A^*\big(AA^*+\omega\tilde{I}\,\big)^{-1}.$$

Таким образом, мы можем переписать (2.19) как

$$A^- y = \lim_{\omega\to+0} \overline{\big(A^*A+\omega\tilde{I}\,\big)^{-1}A^*}\, y \tag{2.25}$$

для любого $y \in \mathcal{D}(A^-)$.

Будем писать $A \overset{\mathcal{M}}{=} B$, если равенство $Ax = Bx$ выполнено на всех элементах $x \in \mathcal{M}$ при $\mathcal{M} \subset \mathcal{D}(A)\cap\mathcal{D}(B)$, чтобы отличать его от строгого операторного равенства $A = B$, подразумевающего совпадение $\mathcal{D}(A) = \mathcal{D}(B)$ областей определения. В таких обозначениях равенства (2.19) и (2.25) могут быть записаны как

$$A^- \overset{\mathcal{D}(A^-)}{=} \lim_{n\to\infty} A^*(AA^*+\omega_n I)^{-1}, \qquad A^- \overset{\mathcal{D}(A^-)}{=} \lim_{\omega\to+0} \overline{\big(A^*A+\omega\tilde{I}\,\big)^{-1}A^*}\, y,$$

где операторы в правых частях соотношений определены на всём $\widetilde{\mathcal{R}}$, а оператор A^- в левых частях — на $\mathcal{D}(A^-)$, т. е. на, вообще говоря, более узком чем $\widetilde{\mathcal{R}}$ множестве.

2.4. Решение линейных уравнений с помощью псевдообращения.

Задача 3. Пусть заданы оператор $A \in \mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ и элемент $y \in \widetilde{\mathcal{R}}$. Требуется найти все элементы $x \in \mathcal{D}(A)$, при которых $Ax = y$.

Решение этой задачи мы, можно сказать, уже получили.

Уравнение $Ax = y$ разрешимо тогда и только тогда, когда $y \in \mathcal{R}(A)$ — это, по сути дела, определение множества $\mathcal{R}(A)$. Если $y \in \mathcal{R}(A)$, то $AA^-y = y$ в силу замечания после свойства 4 (см. также вторую из таблиц, приведённых после свойства 4). Последнее равенство эквивалентно $(\tilde{I} - AA^-)y = 0$. Наоборот, если $AA^-y = y$, то $y \in \mathcal{R}(A)$ (элемент y получен в результате действия A на A^-y), поэтому уравнение $Ax = y$ разрешимо, и A^-y — одно из его решений.

Пусть условие разрешимости выполнены, и мы имеем решение A^-y. Любое другое решение x удовлетворяет уравнению $Ax = y$ и $z = x - A^-y \in \mathcal{N}(A)$. Наоборот, если $x = A^-y + z$, где $z \in \mathcal{N}(A)$, то элемент x удовлетворяет уравнению $Ax = y$. Включение $z \in \mathcal{N}(A)$ имеет место тогда и только тогда, когда $\Pi_0 z = z$, где Π_0 — ортогональный проектор на $\mathcal{N}(A)$. Воспользуемся явным видом проектора для $z \in \mathcal{D}(A)$ (см. вторую таблицу после доказательства свойства 4), получим, что $z \in \mathcal{N}(A)$ тогда и только тогда, когда $z = (I - A^-A)z$. При этом элемент

$A^{-}y \in \mathcal{R}(A^{-}) = \mathcal{N}^{\perp}(A) \cap \mathcal{D}(A)$ ортогонален $z = (I - A^{-}A)z \in \mathcal{N}(A)$, поэтому $\|x\|^2 = \|A^{-}y\|^2 + \|(I - A^{-}A)z\|^2$.

Решение задачи 3 таково. Уравнение $Ax = y$ разрешимо, если и только если

$$(\tilde{I} - AA^{-})y = 0, \quad y \in \mathcal{D}(A^{-}); \tag{2.26}$$

множество всех её решений можно записать как

$$\{x = A^{-}y + (I - A^{-}A)z, \quad z \in \mathcal{D}(A)\}, \tag{2.27}$$

причём слагаемые в последнем равенстве ортогональны, и $A^{-}y$ есть (единственное) решение, имеющее минимальную норму.

Перейдём теперь к операторным уравнениям.

Задача 4. Пусть заданы операторы $A \in \mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ и $Y \in \mathbb{L}(\widehat{\mathcal{R}} \mapsto \widetilde{\mathcal{R}})$. Требуется найти линейный оператор $X \in \mathbb{L}(\widehat{\mathcal{R}} \mapsto \mathcal{R})$ такой, что при всех $u \in \mathcal{D}(Y)$ справедливо равенство $AXu = Yu$.

Напомним, мы можем записать решаемое уравнение как $AX \overset{\mathcal{D}(Y)}{=} Y$.

Пусть найдётся линейный оператор X, $\mathcal{D}(X) \subset \mathcal{D}(Y)$, такой, что для любого $u \in \mathcal{D}(Y)$ существует элемент $x = Xu$, при котором $Ax = Yu$. Следовательно, для $y = Yu$ выполнено условие разрешимости (2.26) уравнения $Ax = y$. Перепишем условие (2.26) как $y = AA^{-}y$ или, в исходных обозначениях, как $Yu = AA^{-}Yu$. Заметим, что последнее соотношение эквивалентно равенству $Y = AA^{-}Y$ в операторах.

Пусть теперь $Yu = AA^{-}Yu$ для любого $u \in \mathcal{D}(Y)$. Тогда оператор $X = A^{-}Y$ является решением задачи 2.

Пусть условие $Y = AA^{-}Y$ выполнено, и X — любое решение задачи 2. Положим $Z = X - A^{-}Y$. Тогда для любого $u \in \mathcal{D}(Y)$ мы имеем $Zu \in \mathcal{N}(A)$ или, эквивалентно, $Zu = (I - A^{-}A)Zu$. Наоборот, если $Xu = A^{-}Yu + (I - A^{-}A)Zu$, где Z — произвольный линейный оператор, область определения которого содержит $\mathcal{D}(Y)$, а пространство значений содержится в $\mathcal{D}(A)$, то X_Y — решение задачи 2. Подводим итог.

решение задачи 4. Уравнение $AX \overset{\mathcal{D}(Y)}{=} Y$ разрешимо тогда и только тогда, когда

$$Y = AA^{-}Y,$$

и множество решений задачи 4 описывается формулой

$$\{X \overset{\mathcal{D}(Y)}{=} A^{-}Y + (I - A^{-}A)Z, \quad Z \in \mathbb{L}(\widehat{\mathcal{R}} \mapsto \mathcal{R})\colon\ \mathcal{D}(Z) \supset \mathcal{D}(Y),\ \mathcal{R}(Z) \subset \mathcal{D}(A)\}.$$

Задача 5. Пусть заданы оператор $A \in \mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ и оператор $U \in \mathbb{L}(\mathcal{R} \mapsto \widehat{\mathcal{R}})$ такой, что $\mathcal{D}(U) \supset \mathcal{D}(A)$. Требуется найти оператор $R \in \mathbb{L}(\widetilde{\mathcal{R}} \mapsto \widehat{\mathcal{R}})$, для которого $RAx = Ux$ при всех $x \in \mathcal{D}(A)$.

Пусть уравнение

$$RA \overset{\mathcal{D}(A)}{=} U \tag{2.28}$$

разрешимо. Тогда в силу равенства $A = AA^-A$ (см. свойство 3 псевдообратного оператора) для любого $x \in \mathcal{D}(A)$ мы имеем равенство $Ux = RAx = RAA^-Ax$. По условию $RAz = Uz$ для всякого $z \in \mathcal{D}(A)$. В частности, при любом $x \in \mathcal{D}(A)$ для элемента $z = A^-Ax \in \mathcal{R}(A^-) \subset \mathcal{D}(A)$ мы можем написать равенство $RAz = Uz$ или, возвращаясь к элементу x, — равенство $Ux = UA^-Ax$ для любого $x \in \mathcal{D}(A)$.

Наоборот, если $Ux = UA^-Ax$ для любого $x \in \mathcal{D}(A)$, то оператор UA^- есть решение уравнения (2.28).

Пусть $UA^-Ax = Ux$ для всех $x \in \mathcal{D}(A)$ и, помимо UA^-, взято ещё какое-то решение R. Тогда для оператора $W = R - UA^-$ мы имеем $WAx = Ux - Ux = 0$ для всех $x \in \mathcal{D}(A)$, другими словами, $Wy = 0$ и $Ry = UA^-y$ для всех $y \in \mathcal{R}(A)$. Подведём итог.

решение задачи 3. Уравнение $RA \overset{\mathcal{D}(A)}{=} U$ разрешимо тогда и только тогда, когда $U \overset{\mathcal{D}(A)}{=} UA^-A$, и на множестве $\mathcal{R}(A)$ оператор R, являющийся решением уравнения, определяется однозначно: $R \overset{\mathcal{R}(A)}{=} UA^-$. Действие оператора R на $\mathcal{R}^\perp(A)$ уравнением не фиксируется и может быть задано произвольным образом или не задано вовсе.

Далее нам потребуется описать класс операторов, являющихся решением уравнения (2.28) и заданных на всём пространстве $\widetilde{\mathcal{R}}$ или хотя бы на всюду плотном линейном многообразии $\mathcal{D}(A^-)$. Нетрудно понять, что множество таких решений описывается формулой

$$R = UA^- + W(I - AA^-), \qquad W \in \mathbb{L}(\widetilde{\mathcal{R}} \mapsto \widehat{\mathcal{R}})\colon\ \mathcal{D}(W) \supset \mathcal{D}(A^-). \tag{2.29}$$

В самом деле, согласно общим свойствам псевдообращения $\tilde{I} - AA^-$ — это проектор в $\mathcal{D}(A^-)$ на линейное многообразие $\mathcal{R}^\perp(A) = \mathcal{N}(A^*) = \mathcal{N}(A^-)$, следовательно,

$$W(\tilde{I} - AA^-)y = \begin{cases} 0, & y \in \mathcal{R}(A), \\ Wy, & y \in \mathcal{R}^\perp(A); \end{cases} \qquad Ry = \begin{cases} UA^-y, & y \in \mathcal{R}(A), \\ Wy, & y \in \mathcal{R}^\perp(A). \end{cases}$$

Таким образом, при $y \in \mathcal{R}^\perp(A)$ мы имеем равенство $Ry = Wy$, где W — произвольный линейный оператор (с надлежащей областью определения).

2.5. Задача редукции измерений. Пусть результат измерения описывается элементом $\xi \in \widetilde{\mathcal{R}}$, где $\widetilde{\mathcal{R}}$ — гильбертово пространство. Зададим модель измерения: предположим, что

$$\begin{gathered} \xi = Af + \nu, \qquad A \in \mathbb{L}(\mathcal{R} \mapsto \widetilde{\mathcal{R}}), \\ \nu \text{ — случайный элемент в } \widetilde{\mathcal{R}}, \quad E\nu = 0, \quad \Sigma_\nu = \Sigma. \end{gathered} \tag{2.30}$$

Мы считаем, что неслучайный элемент $f \in \mathcal{R}$ фиксирован, но неизвестен и нет никакой априорной информации о том, какие значения он может принимать (f — произвольный элемент из $\mathcal{R}$). Будем говорить, что условия (2.30) задают линейную *модель* $[A, \Sigma]$ *измерения сигнала* f *без априорной информации* об измеряемом сигнале. Оператор A моделирует измерительный прибор, ν — шум измерения.

Далее мы будем полагать, что $A \in \mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, $\Sigma \in \mathbb{B}(\widetilde{\mathcal{R}} \mapsto \widetilde{\mathcal{R}})$. Кроме того, в силу общих свойств ковариационного оператора $\Sigma = \Sigma^* \geqslant 0$. Вспомним, что если $y \in \mathcal{N}(\Sigma)$, то случайная величина (ν, y) равна нулю с вероятностью единица, другими словами, $(\xi, y) = (Af, y)$ с вероятностью единица. Из этого равенства можно получить точную информацию о сигнале f, таким образом, мы уже не можем говорить о модели без информации об измеряемом сигнале. В силу этих рассуждений будем считать, что оператор $\Sigma > 0$, т. е. невырожден.

Поставим задачу линейного оценивания сигнала Uf: будем считать, что задан линейный оператор U, $\mathcal{D}(U) = \mathcal{R}$, со значениями в (действительном бесконечномерном) гильбертовом пространстве $\widehat{\mathcal{R}}$, и будем искать наилучшую оценку элемента Uf в классе всех оценок, заданных как линейное преобразование результата измерения.

Пусть $R\xi = RAf + R\nu$ — линейное преобразование элемента ξ. Определим ошибку оценивания как максимальную среднеквадратичную погрешность

$$\sup_{f \in \mathcal{R}} \mathrm{E}\|R\xi - Uf\|^2,$$

где наличие точной верхней грани говорит о том, что мы ищем «наихудшую» погрешность среди всех возможных в измерении сигналов f (напомним, элемент f априори произволен). Преобразуем погрешность, учитывая, что $\mathrm{E}\nu = 0$ и, следовательно, $\mathrm{E}R\nu = 0$:

$$\begin{aligned}\sup_{f\in\mathcal{R}} \mathrm{E}\|R\xi - Uf\|^2 &= \sup_{f\in\mathcal{R}} \mathrm{E}\|(RA-U)f + R\nu\|^2 = \\ &= \sup_{f\in\mathcal{R}} \mathrm{E}\big\{\|(RA-U)f\|^2 + 2\mathrm{E}((RA-U)f, R\nu) + \mathrm{E}\|R\nu\|^2\big\} = \\ &= \sup_{f\in\mathcal{R}} \|(RA-U)f\|^2 + 2(R^*(RA-U)f, \mathrm{E}\nu) + \mathrm{E}\|R\nu\|^2 = \\ &= \mathrm{E}\|R\nu\|^2 + \sup_{f\in\mathcal{R}} \|(RA-U)f\|^2 = \\ &= \mathrm{E}\|R\nu\|^2 + \begin{cases} 0, & RA-U=0, \\ +\infty, & RA-U\neq 0. \end{cases}.\end{aligned}$$

Таким образом, погрешность конечна, если выполнены два условия: $RA = U$ и $\mathrm{E}\|R\nu\|^2 < \infty$. В соответствии с этим зададим класс операторов

$$\mathbb{R}_U = \big\{R \in \mathbb{L}(\widetilde{\mathcal{R}} \mapsto \widehat{\mathcal{R}}):\ RA = U,\ \mathrm{E}\|R\nu\|^2 < \infty,\ \mathcal{D}(R) = \widetilde{\mathcal{R}}\big\}. \tag{2.31}$$

Задача редукции в модели $[A, \Sigma]$. Требуется найти оператор $R \in \mathbb{R}_U$, доставляющий следующий минимум:

$$\min_{R\in\mathbb{R}_U} \mathrm{E}\|R\xi - Uf\|^2. \tag{2.32}$$

Оценка $R\xi$ при этом называется *результатом редукции* измерения ξ к виду Uf в модели $[A, \Sigma]$.

Заметим, условие $\sup_{f\in\mathcal{R}} \mathrm{E}\|R\xi - Uf\|^2 < \infty$ равносильно тому, что $ER\xi = Uf$, т. е. $R\xi$ — несмещенная оценка элемента Uf, для любого оператора $R \in \mathbb{R}_U$.

Мы будем искать решение задачи редукции в предположении, что $\Sigma = \tilde{I}$. Это несколько сужает общность рассуждений, но при этом существенно упрощает формулы. Заметим, что если $\Sigma \neq \tilde{I}$, но элементы $\xi, Af, \nu \in \mathcal{D}(\Sigma^{-1/2}) = \mathcal{R}(\Sigma^{1/2})$ (напомним, в силу $\mathcal{N}(\Sigma) = \{0\}$ линейное многообразие $\mathcal{R}(\Sigma) = \mathcal{N}^{\perp}(\Sigma)$ всюду плотно), то мы можем преобразовать результат измерения: $\Sigma^{-1/2}\xi = \Sigma^{-1/2}Af + \Sigma^{-1/2}\nu$, где случайный элемент $\Sigma^{-1/2}\nu$ имеет единичный ковариационный оператор (представляет собой «белый шум»). Переобозначая элемент $\Sigma^{-1/2}\xi$ как ξ, элемент $\Sigma^{-1/2}Af$ — как Af и элемент $\Sigma^{-1/2}\nu$ — как ν, мы приходим к модели $[A, \tilde{I}\,]$.

Рассмотрим уравнение $RA = U$. Как было показано в предыдущем разделе, данное уравнение имеет решение тогда и только тогда, когда $U = UA^-A$, и общий вид решения таков:

$$R = UA^- + W(\tilde{I} - AA^-), \tag{2.33}$$

где $W \in \mathbb{L}(\widetilde{\mathcal{R}} \mapsto \widehat{\mathcal{R}})$ — произвольный линейный оператор с $\mathcal{D}(W) \supset \mathcal{D}(A^-)$. Такой оператор R определен на плотном подмножестве $\mathcal{D}(A^-) = \mathcal{R}(A) \oplus \mathcal{R}^{\perp}(A)$.

Теперь преобразуем $\mathrm{E}\|R\nu\|^2$. Согласно формуле (1.12) и при нашем допущении $\Sigma = \tilde{I}$ мы имеем

$$\mathrm{E}\|R\nu\|^2 = \|R\Sigma^{1/2}\|_2^2 = \|R\|_2^2 = \sum_{i=1}^{\infty} \|R\tilde{e}_i\|^2, \tag{2.34}$$

где $\{\tilde{e}_i\}$ — ОНБ в пространстве $\widetilde{\mathcal{R}}$, который может быть выбран произвольно. Мы возьмём в качестве ОНБ $\{\tilde{e}_k^{\perp}\} \oplus \{\tilde{e}_j^0\}$, где $\{\tilde{e}_k^{\perp}\}$ — ОНБ в линейном многообразии $\mathcal{R}(A)$ и $\{\tilde{e}_j^0\}$ — ОНБ в линейном подпространстве $\mathcal{R}^{\perp}(A)$. Тогда для оператора (2.33) мы имеем

$$\begin{aligned} R\tilde{e}_k^{\perp} &= \big(UA^- + W(\tilde{I} - AA^-)\big)\tilde{e}_k^{\perp} = UA^-\tilde{e}_k^{\perp}, && \tilde{e}_k^{\perp} \in \mathcal{R}(A), \\ R\tilde{e}_j^{0} &= \big(UA^- + W(\tilde{I} - AA^-)\big)\tilde{e}_j^{0} = W(\tilde{I} - AA^-)\big)\tilde{e}_j^{0} = W\tilde{e}_j^{0}, && \tilde{e}_j^{0} \in \mathcal{R}^{\perp}(A), \end{aligned}$$

поскольку оператор $\tilde{I} - AA^-$ есть ортогональный проектор в $\mathcal{D}(A^-)$ на подпространство $\mathcal{R}^{\perp}(A) = \mathcal{N}(A^*) = \mathcal{N}(A^-)$. Таким образом,

$$\mathrm{E}\|R\nu\|^2 = \sum_{k=1}^{\infty}\|R\tilde{e}_k^{\perp}\|^2 + \sum_{j=1}^{\infty}\|R\tilde{e}_j^{0}\|^2 = \sum_{k=1}^{\infty}\|UA^-\tilde{e}_k^{\perp}\|^2 + \sum_{j=1}^{\infty}\|W\tilde{e}_j^{0}\|^2 \geqslant \sum_{k=1}^{\infty}\|UA^-\tilde{e}_k^{\perp}\|^2,$$

и мы получаем, что минимум $\mathrm{E}\|R\nu\|^2$ среди всех операторов вида (2.33) достигается при $W = 0$.

Оператор UA^- определён на $\mathcal{D}(A^-)$, но условие $\mathrm{E}\|R\nu\|^2 = \sum_{k=1}^{\infty}\|UA^-\tilde{e}_k^{\perp}\|^2 < \infty$ гарантирует, что этот оператор ограничен (более того, на $\mathcal{D}(A^-)$ он является оператором Гильберта–Шмидта), поэтому его можно доопределить по непрерывности (замкнуть) на всём пространстве $\widetilde{\mathcal{R}} = \overline{\mathcal{D}(A)}$, и замыкание этого оператора $\overline{UA^-}$ принадлежит пространству $\mathbb{H}(\widetilde{\mathcal{R}} \mapsto \widehat{\mathcal{R}})$. Итак, решение задачи редукции даётся следующей теоремой.

Теорема 6.

Задача (2.32) редукции в модели $[A, \tilde{I}\,]$ разрешима, если и только если выполнены условия $U = UA^-A$ и $\overline{UA^-} \in \mathbb{H}(\widetilde{\mathcal{R}} \mapsto \widehat{\mathcal{R}})$, и наилучшая в среднеквадратичном несмещенная линейная оценка элемента Uf есть $\overline{UA^-}\,\xi$; погрешность такой оценки равна

$$\|\overline{UA^-}\,\|_2^2 = \sum_k \|UA^-\tilde{e}_k^{\perp}\|^2, \tag{2.35}$$

где $\{\tilde{e}_k^{\perp}\}$ — произвольный ОНБ в $\mathcal{R}(A)$.

ГЛАВА 3
ЭФФЕКТИВНЫЙ РАНГ МОДЕЛИ $[A, \Sigma]$

3.1. Сингулярные базисы псевдообратного оператора. Для построения эффективного ранга мы будем использовать математический аппарат теории сингулярных чисел, изложенный в приложении C. Далее будем считать, что оператор A всюду определён и компактен. Как обычно, положим $r = \operatorname{rank} A \dim \mathcal{R}(A)$ и будем рассматривать как случай $r < \infty$, так и случай $r = \infty$, уделяя последнему основное внимание. Введём для оператора A (см. формулы (C.27) приложения C) его сингулярные базисы $\{e_k\} \oplus \{e_i^0\}$ и $\{\tilde{e}_k\} \oplus \{\tilde{e}_j^0\}$ и строго положительные сингулярные числа $a_1 \geqslant a_2 \geqslant \cdots$:

$$\begin{aligned} &Ae_k = a_k \tilde{e}_k, && A^* \tilde{e}_k = a_k e_k, && k = 1, \ldots, r = \operatorname{rank} A \leqslant \infty, \\ &Ae_j^0 = 0, && A^* \tilde{e}_i^0 = 0. && \end{aligned} \tag{3.1}$$

Напомним, что первая пара равенств устанавливает взаимно однозначное соответствие между ОНБ $\{e_k\}$ и $\{\tilde{e}_k\}$, таким образом, эти базисы содержат «одинаковое» количество элементов, равное рангу оператора A, даже в случае, когда $r = \infty$. Что касается равенств в второй строке, то в них индексы пробегают, вообще говоря, разные значения: $Ae_j^0 = 0$ для $j = 1, \ldots, \dim \mathcal{N}(A)$ и $A^* \tilde{e}_i = 0$ для $i = 1, \ldots, \dim \mathcal{N}(A^*)$, однако эти равенства не имеют для нас большого значения, поэтому мы не пишем в них, как меняются индексы.

Из (3.1) имеем $A^- \tilde{e}_k = a_k^{-1} A^- A e_k = a_k^{-1} e_k$, потому что $e_k \in \mathcal{R}(A^*)$, а оператор $A^- A$ проецирует на $\mathcal{R}(A^*)$. Далее, $A^- \tilde{e}_j^0 = 0$, поскольку $\tilde{e}_j^0 \in \mathcal{N}(A^*) = \mathcal{N}(A^-)$. Аналогичные формулы можно написать и для $(A^-)^* = (A^*)^-$. Таким образом,

$$\begin{aligned} &A^- \tilde{e}_k = \frac{1}{a_k} e_k, && (A^-)^* e_k = \frac{1}{a_k} e_k, && k = 1, \ldots, r, \\ &A^- \tilde{e}_j^0 = 0, && (A^-)^* e_j^0 = 0, && \end{aligned} \tag{3.2}$$

и мы получаем равенства, совершенно идентичные (3.1), с той лишь разницей, что теперь сингулярные числа упорядочены по возрастанию: $1/a_k \leqslant 1/a_{k+1}$.

Положим для краткости $\mathcal{L}_m = \mathcal{L}(e_1, \ldots, e_m)$ и посмотрим, во что перейдут равенства (см. формулы (C.30) из приложения C)

$$a_{m+1} = \sup_{\substack{x \in \mathcal{L}_m^\perp, \\ \|x\| \neq 0}} \frac{\|Ax\|}{\|x\|}, \qquad m = 0, 1, \ldots, r = \operatorname{rank} A \leqslant \infty, \tag{3.3}$$

где мы для единообразия положили по определению $\mathcal{L}_1^\perp = \mathcal{L}^\perp(e_1, \ldots, e_m) = \mathcal{R}$ при $m = 0$ (тогда $a_1 = \|A\|$). Кроме того, в случае $r < \infty$ мы считаем, что $a_{r+1} = 0$ при $m = r$ в силу $\mathcal{L}^\perp(e_1, \ldots, e_r) = \mathcal{N}(A)$.

Пусть $m = 0$. Разложим произвольный элемент $x \in \mathcal{R}$ на ортогональные составляющие $x^\perp \in \mathcal{N}^\perp(A)$ и $x^0 \in \mathcal{N}(A)$ как $x = x^\perp + x^0$ и преобразуем формулу (3.3) при $m = 0$, т. е. для первого сингулярного числа (здесь и далее мы не пишем включения $x \in \mathcal{R}$ или $y \in \widetilde{\mathcal{R}}$ под знаком точной верхней грани, поскольку они не ограничивают область):

$$a_1^2 = \sup_{\|x\| \neq 0} \frac{\|A(x^\perp + x^0)\|^2}{\|x^\perp\|^2 + \|x^0\|^2} = \sup_{\|x\| \neq 0} \frac{\|Ax^\perp\|^2}{\|x^\perp\|^2 + \|x^0\|^2} = \sup_{\substack{x^\perp \in \mathcal{N}^\perp(A), \\ \|x^\perp\| \neq 0}} \frac{\|Ax^\perp\|^2}{\|x^\perp\|^2}, \tag{3.4}$$

где мы воспользовались тем, что наличие слагаемого $\|x^0\|^2 > 0$ в знаменателе только уменьшает значение дроби. Поскольку мы предположили, что $A \in \mathbb{CO}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, т. е. в частности, $\mathcal{D}(A) = \mathcal{R}$, можно написать $\mathcal{R}(A^-) = \mathcal{N}^\perp(A) \cap \mathcal{D}(A) = \mathcal{N}^\perp(A)$, отсюда для любого $x^\perp \in \mathcal{N}^\perp(A)$ найдётся $y \in \mathcal{D}(A^-)$ такой, что $x^\perp = A^- y$. Разложим и этот элемент на ортогональные составляющие, $y = y^\perp + y^0$, в соответствии с разложением $\mathcal{D}(A^-) = \mathcal{R}(A) \oplus \mathcal{N}(A^*)$. Тогда

$$Ax^\perp = AA^- y = AA^-(y^\perp + y^0) = y^\perp, \qquad x^\perp = A^- y = A^-(y^\perp + y^0) = A^- y^\perp,$$

в силу того что оператор AA^- — ортогональный проектор в $\mathcal{D}(A^-)$ на линейное многообразие $\mathcal{R}(A)$ и $\mathcal{N}(A^*) = \mathcal{N}(A^-)$. Причём $x^\perp \neq 0$ тогда и только тогда, когда $y^\perp \neq 0$, поскольку операторы A и A^- осуществляют взаимно однозначное соответствие между $\mathcal{N}^\perp(A)$ и $\mathcal{R}(A)$. Таким образом, мы можем переписать правую часть равенства (3.4) и получаем

$$a_1^2 = \sup_{\substack{y^\perp \in \mathcal{R}(A), \\ \|y^\perp\| \neq 0}} \frac{\|y^\perp\|^2}{\|A^- y^\perp\|^2} = \left(\inf_{\substack{y^\perp \in \mathcal{R}(A), \\ \|y^\perp\| \neq 0}} \frac{\|A^- y^\perp\|^2}{\|y^\perp\|^2} \right)^{-1}$$

или, заменив для краткости записи $y^\perp$ на y,

$$\frac{1}{a_1^2} = \inf_{\substack{y \in \mathcal{R}(A), \\ \|y\| \neq 0}} \frac{\|A^- y\|^2}{\|y\|^2}. \tag{3.5}$$

Пусть теперь $1 < m < r \leqslant \infty$. Сузим пространство $\mathcal{R}$ до $\mathcal{L}_m^\perp = \mathcal{L}^\perp(e_1, \ldots, e_m)$ и перейдем к сужению A_m оператора A на $\mathcal{L}_m^\perp$, другими словами, исключим из пространства $\mathcal{R}$ первые m «размерностей». Очевидно, что формулы (3.1) и (3.2) сохраняют свою силу:

$$
\begin{aligned}
&A_m e_k = A e_k = a_k \tilde{e}_k, \qquad A_m^- \tilde{e}_k = A_- \tilde{e}_k = \frac{1}{a_k} e_k, \qquad && k = m+1, \ldots, r,\\
&A_m e_j^0 = A e_j^0 = 0, && j = 1, \ldots, \dim \mathcal{N}(A)
\end{aligned}
$$

Следовательно,

$$
\mathcal{R}(A_m) = \mathcal{L}^\perp(\tilde{e}_1, \ldots, \tilde{e}_m) \cap \mathcal{R}(A) = \mathcal{L}(\tilde{e}_{m+1}, \ldots, \tilde{e}_r), \qquad \mathcal{N}(A_m) = \mathcal{N}(A).
$$

По аналогии с равенством (3.5) получаем

$$
\frac{1}{a_{m+1}^2} = \inf_{\substack{y \in \widetilde{\mathcal{L}}_m^\perp \cap \mathcal{R}(A),\\ \|y\| \neq 0}} \frac{\|A_m^- y\|^2}{\|y\|^2}.
$$

Поскольку $A_m^- y = A^- y$ для $y \in \mathcal{L}(\tilde{e}_{m+1}, \ldots, \tilde{e}_r)$, мы имеем формулу типа (3.5) в случае $1 < m < r = \operatorname{rank} A$:

$$
\frac{1}{a_{m+1}^2} = \inf_{\substack{y \in \widetilde{\mathcal{L}}_m^\perp \cap \mathcal{R}(A),\\ \|y\| \neq 0}} \frac{\|A^- y\|^2}{\|y\|^2}, \qquad \widetilde{\mathcal{L}}_m^\perp = \mathcal{L}(\tilde{e}_{m+1}, \ldots, \tilde{e}_r), \qquad 1 \leqslant m < r. \tag{3.6}
$$

Поскольку $\mathcal{L}(\tilde{e}_1, \ldots, \tilde{e}_r) = \mathcal{R}(A)$ для $m = 0$, мы можем считать, что формула (3.6) верна и в случае $m = 1$. Если $r < \infty$, то $a_{r+1} = 0$, и тогда формула (3.6) не имеет смысла при $m \geqslant r$.

Докажем основное для нахождения эффективного ранга свойство сингулярных базисов.

Теорема 7.
Пусть натуральное число $m < r = \operatorname{rank} A$ и $\tilde{e}_1, \ldots, \tilde{e}_m$ – первые m элементов сингулярного базиса оператора A. Тогда для любого конечного натурального числа $n > m$ и для любой ортонормированной системы элементов $\widetilde{\varphi}_1, \ldots, \widetilde{\varphi}_n$ в $\mathcal{R}(\mathcal{A})$ справедлива следующая оценка:

$$
\sum_{i=1}^n \|A^- \widetilde{\varphi}_i\|^2 \geqslant \sum_{k=1}^m \frac{1}{a_k^2} + \sum_{k=m+1}^n \frac{1}{a_{m+1}^2}. \tag{3.7}
$$

Доказательство. Разложим элемент $\widetilde{\varphi}_i \in \mathcal{R}(A)$ по сингулярным базисам и вместе с ним — элемент $A^- \widetilde{\varphi}_i$:

$$
\widetilde{\varphi}_i = \sum_{k=1}^m (\widetilde{\varphi}_i, \tilde{e}_k) \tilde{e}_k + \widetilde{\varphi}_i^\perp, \qquad A^- \widetilde{\varphi}_i = \sum_{k=1}^m \frac{1}{a_k} (\widetilde{\varphi}_i, \tilde{e}_k) e_k + A^- \widetilde{\varphi}_i^\perp, \tag{3.8}
$$

где $\widetilde{\varphi}_i^\perp \in \mathcal{L}^\perp(\tilde{e}_1, \ldots, \tilde{e}_m)$ Тогда для каждого $k = 1, \ldots, m$

$$\begin{aligned} (\widetilde{\varphi}_i^\perp, \tilde{e}_k) &= (\widetilde{\varphi}_i, \tilde{e}_k) - (\widetilde{\varphi}_i, \tilde{e}_k) = 0, \\ (A^- \widetilde{\varphi}_i^\perp, e_k) &= (\widetilde{\varphi}_i^\perp, (A^-)^* e_k) = (\widetilde{\varphi}_i^\perp, (A^*)^- e_k) = \frac{1}{a_k}(\widetilde{\varphi}_i^\perp, \tilde{e}_k) = 0. \end{aligned} \tag{3.9}$$

Отсюда следует, что в каждом из равенств (3.8) слагаемые ортогональны, поэтому

$$1 = \|\widetilde{\varphi}_i\|^2 = \sum_{k=1}^m (\widetilde{\varphi}_i, \tilde{e}_k)^2 + \|\widetilde{\varphi}_i^\perp\|^2, \qquad \|A^- \widetilde{\varphi}_i\|^2 = \sum_{k=1}^m \frac{1}{a_k^2}(\widetilde{\varphi}_i, \tilde{e}_k)^2 + \|A^- \widetilde{\varphi}_i^\perp\|^2. \tag{3.10}$$

Кроме того, вследствие первого из соотношений (3.9) и условия $\widetilde{\varphi}_i \in \mathcal{R}(A)$ мы имеем $\widetilde{\varphi}_i^\perp \in \widetilde{\mathcal{L}}_m^\perp \cap \mathcal{R}(A) = \mathcal{L}(\tilde{e}_{m+1}, \ldots, \tilde{e}_r)$. Таким образом, применив формулу (3.6), мы можем записать оценку

$$\frac{\|A^- \widetilde{\varphi}_i^\perp\|^2}{\|\widetilde{\varphi}_i^\perp\|^2} \geqslant \inf_{\substack{y \in \mathcal{L}^\perp(\tilde{e}_{m+1}, \ldots, \tilde{e}_r), \\ \|y\| \neq 0}} \frac{\|A^- y\|^2}{\|y\|^2} = \frac{1}{a_{m+1}^2}$$

или, эквивалентно,

$$\|A^- \widetilde{\varphi}_i^\perp\|^2 \geqslant \frac{1}{a_{m+1}^2} \|\widetilde{\varphi}_i^\perp\|^2 = \frac{1}{a_{m+1}^2} \left(1 - \sum_{k=1}^m (\widetilde{\varphi}_i, \tilde{e}_k)^2\right),$$

где мы учли первое из равенств (3.10).

Теперь с учётом получившейся оценки найдём разность Δ между левой и правой частями доказываемого неравенства (3.7) и покажем, что $\Delta \geqslant 0$. Имеем

$$\begin{aligned} \Delta &= \sum_{i=1}^n \|A^- \widetilde{\varphi}_i\|^2 - \sum_{k=1}^m \frac{1}{a_k^2} - \sum_{k=m+1}^n \frac{1}{a_{m+1}^2} = \\ &= \sum_{i=1}^n \left(\sum_{k=1}^m \frac{1}{a_k^2} (\widetilde{\varphi}_i, \tilde{e}_k)^2 + \|A^- \widetilde{\varphi}_i^\perp\|^2 \right) - \sum_{k=1}^m \frac{1}{a_k^2} - \sum_{k=m+1}^n \frac{1}{a_{m+1}^2} \geqslant \\ &\geqslant \sum_{i=1}^n \sum_{k=1}^m \frac{1}{a_k^2} (\widetilde{\varphi}_i, \tilde{e}_k)^2 + \sum_{i=1}^n \frac{1}{a_{m+1}^2} \left(1 - \sum_{k=1}^m (\widetilde{\varphi}_i, \tilde{e}_k)^2\right) - \sum_{k=1}^m \frac{1}{a_k^2} - \sum_{k=m+1}^n \frac{1}{a_{m+1}^2}. \end{aligned} \tag{3.11}$$

Рассмотрим сначала члены, содержащие множители $1/a_k^2$ с $k = 1, \ldots, m$:

$$\Delta_1 = \sum_{i=1}^n \sum_{k=1}^m \frac{1}{a_k^2} (\widetilde{\varphi}_i, \tilde{e}_k)^2 - \sum_{k=1}^m \frac{1}{a_k^2} = -\sum_{k=1}^m \frac{1}{a_k^2} \left(1 - \sum_{i=1}^n (\widetilde{\varphi}_i, \tilde{e}_k)^2\right).$$

Заметим, что с учётом нормировки $\|\tilde{e}_k\|^2 = 1$, введя Π_Φ — ортогональный проектор на $\mathcal{L}(\widetilde{\varphi}_1, \ldots, \widetilde{\varphi}_n)$, т. е. $\widetilde{\Pi}_\Phi y = \sum_{i=1}^n (\widetilde{\varphi}_i, y)\widetilde{\varphi}_i$ для любого $y \in \widetilde{\mathcal{R}}$, мы можем написать

$$1 = \|\tilde{e}_k\|^2 = \|\widetilde{\Pi}_\Phi \tilde{e}_k\|^2 + \|(\tilde{I} - \widetilde{\Pi}_\Phi)\tilde{e}_k\|^2 = \sum_{i=1}^n (\widetilde{\varphi}_i, \tilde{e}_k)^2 + \|(\tilde{I} - \widetilde{\Pi}_\Phi)\tilde{e}_k\|^2.$$

Таким образом, мы можем записать Δ_1 как

$$\Delta_1 = -\sum_{k=1}^{m} \frac{1}{a_k^2} \|(\tilde{I} - \widetilde{\Pi}_\Phi)\tilde{e}_k\|^2. \tag{3.12}$$

Теперь рассмотрим оставшиеся (пропорциональные $1/a_{m+1}^2$) члены в правой части неравенства (3.11), заменив один раз индекс суммирования i на k:

$$\begin{aligned}
\Delta_2 &= \sum_{i=1}^{n} \frac{1}{a_{m+1}^2} \left(1 - \sum_{k=1}^{m} (\widetilde{\varphi}_i, \tilde{e}_k)^2\right) - \sum_{k=m+1}^{n} \frac{1}{a_{m+1}^2} = \\
&= \left(\sum_{k=1}^{n} \frac{1}{a_{m+1}^2} - \sum_{k=m+1}^{n} \frac{1}{a_{m+1}^2}\right) - \sum_{i=1}^{n}\sum_{k=1}^{m} \frac{1}{a_{m+1}^2} (\widetilde{\varphi}_i, \tilde{e}_k)^2 = \\
&= \sum_{k=1}^{m} \frac{1}{a_{m+1}^2} - \sum_{k=1}^{m}\sum_{i=1}^{n} \frac{1}{a_{m+1}^2} (\widetilde{\varphi}_i, \tilde{e}_k)^2 = \sum_{k=1}^{m} \frac{1}{a_{m+1}^2} \left(1 - \sum_{i=1}^{n} (\widetilde{\varphi}_i, \tilde{e}_k)^2\right).
\end{aligned}$$

Мы снова видим в круглых скобках в правой части выражение, которое в соответствии с нашими обозначениями равно $\|(\tilde{I} - \widetilde{\Pi}_\Phi)\tilde{e}_k\|^2$. В результате

$$\Delta_2 = \sum_{k=1}^{m} \frac{1}{a_{m+1}^2} \|(\tilde{I} - \widetilde{\Pi}_\Phi)\tilde{e}_k\|^2. \tag{3.13}$$

Подставляя (3.12) и (3.13) в (3.11), получаем

$$\Delta \geqslant \Delta_1 + \Delta_2 = \sum_{k=1}^{m} \left(\frac{1}{a_{m+1}^2} - \frac{1}{a_k^2}\right) \|(\tilde{I} - \widetilde{\Pi}_\Phi)\tilde{e}_k\|^2 \geqslant 0,$$

поскольку $a_{m+1}^2 \leqslant a_k^2$ при $k \leqslant m$. Таким образом, неравенство (3.7) доказано.

Пусть $n = m + 1 < r$, тогда вторая сумма в (3.7) будет содержать одно слагаемое, и его можно добавить к первой сумме, записав в правой части (3.7) единую сумму по k от $k = 1$ до $k = m + 1 = n$:

$$\sum_{i=1}^{n} \|A^- \widetilde{\varphi}_i\|^2 \geqslant \sum_{k=1}^{n} \frac{1}{a_k^2}.$$

С другой стороны,

$$\sum_{k=1}^{n} \|A^- \tilde{e}_k\|^2 = \sum_{i=1}^{n} \frac{1}{a_i^2}.$$

Таким образом, для любого $n < r = \operatorname{rank} A$ справедливо следующее утверждение:

$$\sum_{k=1}^{n} \frac{1}{a_k^2} = \sum_{i=1}^{n} \|A^- \tilde{e}_k\|^2 = \min_{\{\widetilde{\varphi}_k\} \subset \mathcal{R}(A)} \sum_{k=1}^{n} \|A^- \widetilde{\varphi}_k\|^2, \tag{3.14}$$

где минимум берётся по всем ортонормированным системам элементов $\widetilde{\varphi}_1, \ldots, \widetilde{\varphi}_n$, лежащих в $\mathcal{R}(A)$.

3.2. Эффективный ранг модели $[A, \Sigma]$. Погрешность редукции (2.35) определяется моделью измерения, в том числе операторами A и $\Sigma = \tilde{I}$, но она также зависит от оператора U. Если модель измерения фиксирована, единственным способом уменьшить погрешность является изменение оператора U (разумеется, при условии, что мы не подвергаем сомнению истинность само́й модели). Можно попытаться оценить не весь сигнал f, а некоторую его «часть». Для линейных задач естественным способом вырезания части сигнала является операция ортогонального проектирования.

Итак, пусть $U = \Pi$ — ортогональный проектор. Запишем погрешность редукции:

$$h(\Pi) = \|\overline{\Pi A^-}\|_2^2 = \sum_k \|\Pi A^- \tilde{g}_k^\perp\|^2,$$

где $\{\tilde{g}_k^\perp\}$ — произвольный ОНБ в $\mathcal{R}(A)$. При этом необходимо сохранить условие разрешимости $\Pi = \Pi A^- A$. Сформулируем его в эквивалентном, но более удобном для нас виде. А именно, сначала перепишем $\Pi = \Pi A^- A$ как $\Pi(I - A^- A) = 0$, и это означает, что $\mathcal{R}(I - A^- A) \subset \mathcal{N}(\Pi)$. Далее, поскольку $I - A^- A$ представляет собой ортогональный проектор на $\mathcal{N}(A)$, последнее включение эквивалентно $\mathcal{N}(A) \subset \mathcal{N}(\Pi)$. Отметим, что это совершенно естественное условие: если $Af = 0$, то измерение $\xi = Af + \nu = \nu$ представляет собой чистый «шум», и элемент Πf можно по такому измерению оценить, только если $\Pi f = 0$.

Перейдём в $\mathcal{N}(A) \subset \mathcal{N}(\Pi)$ к ортогональным дополнениям, получим ещё одно эквивалентное соотношение $\mathcal{R}(\Pi) \subset \overline{\mathcal{R}(A^*)}$, где мы приняли во внимание, что $\operatorname{rank}\Pi = \dim \mathcal{R}(\Pi) < \infty$, и не поставили символ замыкания над $\mathcal{R}(\Pi)$. Таким образом, условие $\Pi = \Pi A^- A$ можно переписать как $\mathcal{R}(\Pi) \subset \overline{\mathcal{R}(A^*)}$. Отсюда, кстати, следует, что $\operatorname{rank}\Pi \leqslant \operatorname{rank} A = \dim \mathcal{R}(A)$.

Преобразуем выражение для погрешности, введя произвольные ОНБ $\{\tilde{g}_k\}$ в линейном многообразии $\mathcal{D}(A^-)$ и $\{\varphi_i\}$ в линейном подпространстве $\mathcal{R}(\Pi) \subset \overline{\mathcal{R}(A^*)}$; отметим, что ОНБ $\{\varphi_i\}$ содержит конечное число элементов, равное $\operatorname{rank}\Pi < \infty$. Итак,

$$h(\Pi) = \|\overline{\Pi A^-}\|_2^2 = \sum_{k=1}^{\infty} \|\Pi A^- \tilde{g}_k\|^2 = \sum_{k=1}^{\infty} \sum_{i=1}^{\operatorname{rank}\Pi} (\Pi A^- \tilde{g}_k, \varphi_i)^2 = \sum_{i=1}^{\operatorname{rank}\Pi} \sum_{k=1}^{\infty} (A^- \tilde{g}_k, \varphi_i)^2, \tag{3.15}$$

где мы воспользовались тем, что $\operatorname{rank}\Pi$ конечен, и поменяли порядок суммирования, а также учли, что $\Pi^* \varphi_i = \Pi \varphi_i = \varphi_i$.

Покажем, что для выполнения условия $h(\Pi) < \infty$ необходимо более точное, чем $\varphi_i \in \overline{\mathcal{R}(A^*)}$, включение $\varphi_i \in \mathcal{R}(A^*)$ для всех $i = 1, \ldots, \operatorname{rank}\Pi$. Рассмотрим случай, когда $\varphi_i \in \overline{\mathcal{R}(A^*)} \setminus \mathcal{R}(A)$. Тогда $\varphi_i \notin \mathcal{D}(A^*)^- = \mathcal{R}(A^*) \oplus \mathcal{R}^\perp(A^*)$. Но $(A^*)^- = (A^-)^*$, следовательно, $\varphi_i \notin \mathcal{D}(A^-)^*$, и это означает, что

$$\sup_{\substack{y\in\mathcal{D}(A^-),\\ \|y\|=1}} |(A^-y, \varphi_i)| = \infty.$$

Таким образом, мы можем выбрать в (3.15) надлежащим образом один базисный элемент (скажем, $\tilde{g}_1$) и сделать слагаемое $(A^-\tilde{g}_1, \varphi_i)^2$, а вместе с ним и погрешность $h(\Pi) = \|\overline{\Pi A^-}\|_2^2$ больше любого наперед заданного числа. Поскольку норма Гильберта–Шмидта либо конечна для любого базиса $\{\tilde{g}_k\}$ в линейном многообразии $\mathcal{D}(A^-)$, либо бесконечна для любого базиса, мы получаем, что в этом случае $\|\overline{\Pi A^-}\|_2^2 = \infty$. Таким образом, условие $h(\Pi) < \infty$ влечёт $\{\varphi_i\} \subset \mathcal{R}(A)$.

Предположим, что мы хотим оценить конечное, но не меньшее чем n число координат вектора f. В результате мы получаем естественный класс конечномерных ортогональных проекторов, для которых задача редукции в модели $[A, \Sigma]$ разрешима и погрешность $h(\Pi)$ конечна:

$$\mathbb{P}_n = \big\{\Pi\colon\ \mathcal{R}(\Pi) \subset \mathcal{R}(A^*),\ n \leqslant \operatorname{rank}\Pi,\ n < \infty\big\}. \tag{3.16}$$

Поставим следующую вариационную задачу: в указанном множестве требуется найти ортогональный проектор, для которого погрешность минимальна, т. е. найдём проектор, на котором достигается

$$\min_{\Pi\in\mathbb{P}_n} \|\overline{\Pi A^-}\|_2^2. \tag{3.17}$$

При условии $\varphi_i \in \mathcal{R}(A^*) \subset \mathcal{D}(A^-)^*$ можно записать $(A^-\tilde{g}_k, \varphi_i) = (\tilde{g}_k, (A^-)^*\varphi_i)$, и погрешность (3.15) приобретает вид

$$\|\overline{\Pi A^-}\|_2^2 = \sum_{i=1}^{\operatorname{rank}\Pi} \sum_{k=1}^{\infty} (\tilde{g}_k, (A^-)^*\varphi_i)^2 = \sum_{i=1}^{\operatorname{rank}\Pi} \|(A^-)^*\varphi_i\|^2,$$

поскольку $\{\tilde{g}_k\}$ — ОНБ в $\mathcal{D}(A^-)$ и, следовательно, в $\widetilde{\mathcal{R}} = \overline{\mathcal{D}(A^-)}$. Видно, что погрешность тем больше, чем больше $\operatorname{rank}\Pi$, поэтому минимум $\|\overline{\Pi A^-}\|_2^2$ достигается на проекторе из класса $\mathbb{P}_n$, ранг которого минимален, т. е. равен n. Таким образом, мы можем переформулировать задачу (3.17): требуется найти

$$\min_{\{\varphi_i\}\subset\mathcal{R}(A^*)} \sum_{i=1}^{n} \|(A^-)^*\varphi_i\|^2, \tag{3.18}$$

где $\{\varphi_i\}$ — ортонормированная система из n элементов в $\mathcal{R}(A^*)$. Мы видим полную идентичность данного минимума и минимума в (3.14), следует только заменить в (3.14) оператор A на A^*. В результате получаем, что

$$\min_{\pi \in \mathbb{P}_n} \|\overline{\Pi A^-}\|_2^2 = \min_{\{\varphi_i\} \subset \mathcal{R}(A^*)} \sum_{i=1}^{n} \|(A^-)^* \varphi_i\|^2 = \sum_{i=1}^{n} \|(A^-)^* e_i\|^2 = \sum_{i=1}^{n} \frac{1}{a_i^2}, \tag{3.19}$$

а проектор Π_n, доставляющий минимум, проецирует на $\mathcal{L}(e_1, \ldots, e_n)$ — линейную оболочку первых n сингулярных векторов оператора A.

Введём функцию $H\colon \mathbb{N} \mapsto \mathbb{R}_+$ формулой

$$H(n) = \sum_{k=1}^{n} \frac{1}{a_k}, \qquad 1 \leqslant n \leqslant r = \operatorname{rank} A, \quad n < \infty.$$

Эта функция строго возрастает от минимального значения $H_\circ = H(1) = a_1^{-2}$ до максимального значения значения $H^\circ = H(r)$, которое равно бесконечности в случае $r = \infty$. Последнее утверждение вытекает из того, что для компактного оператора A^*A последовательность $\{a_n\}$ его собственных значений стремится к нулю.

Зададим теперь функцию $\rho\colon [H_\circ, H^\circ) \mapsto \mathbb{N}$ равенством

$$\rho(\delta) = \max_{n\colon H(n) \leqslant \delta} n \tag{3.20}$$

и назовем $\rho(\cdot)$ *эффективным рангом модели измерения* $[A, \tilde{I}\,]$. Доопределим её как $\rho(\delta) = 0$ при $0 < \delta < H_\circ$ и как $\rho(\delta) = r$ при $\delta \geqslant H^\circ$, если $r < \infty$.

Значение $\rho(\delta) = n$ говорит о том, что если мы решаем задачу редукции в модели $[A, \tilde{I}\,]$, то с погрешностью, не большей δ, можно оценить не более чем n координат вектора f. Если $r = \operatorname{rank} A < \infty$ и δ достаточно велико, то эффективный ранг совпадает с рангом оператора A.

ПРИЛОЖЕНИЕ А
НЕКОТОРЫЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ВЕРОЯТНОСТЕЙ

В этом приложении мы приведём некоторые определения и теоремы теории вероятностей, которые потребуются нам в основном курсе. Мы не даём строгих обоснований сходимости интегралов, поскольку видим цель этого материала прежде всего в ознакомлении с методами и инструментами теории вероятности. Хорошо известные теоремы приводятся без доказательства, и основное внимание уделено тем утверждениям, в основном из теории меры, которые обычно не доказывают в общих курсах теории вероятностей.

А.1. Аксиомы вероятности. *Вероятностным пространством* называется совокупность трёх математических объектов $[\Omega, \mathcal{F}, \mathrm{P}]$, где:

Ω — множество некоторых элементов, которые называются *элементарными исходами*,

$\mathcal{F}$ — некоторая сигма-алгебра его подмножеств[4)], каждое подмножество $A \in \mathcal{F}$ называется *событием*,

P — вероятность, заданная на сигма-алгебре $\mathcal{F}$.

Вероятностью называется функция $\mathrm{P}(\cdot)$, которая единственным образом сопоставляет любому событию A конечное действительное число $\mathrm{P}(A)$ и удовлетворяет следующим аксиомам:

- неотрицательность вероятности: $\mathrm{P}(A) \geqslant 0$ для любого $A \in \mathcal{F}$,
- счётная аддитивность вероятности:

$$\mathrm{P}(A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots) = \mathrm{P}(A_1) + \mathrm{P}(A_2) + \cdots + \mathrm{P}(A_n) + \cdots,$$

если $A_k \in \mathcal{F}$ и $A_k \cap A_{k'} = \varnothing$ при $k \neq k'$,

- нормировка вероятности: $\mathrm{P}(\Omega) = 1$.

Важнейшим следствием из аксиом является теорема о непрерывности вероятности:

$$\mathrm{P}(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} \mathrm{P}(A_n), \tag{A.1}$$

где предел последовательности событий $\{A_n\}$ определяется следующим образом: вводятся события, которые называются верхним и нижним пределами последовательности $\{A_n\}$,

$$\limsup A_n = \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} A_n, \qquad \liminf A_n = \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} A_n, \tag{A.2}$$

[4)] Множество подмножеств называется сигма-алгеброй, если оно замкнуто относительно операций дополнения, конечного или счётного объединения и конечного или счётного пересечения множеств.

а пределом последовательности $\{A_n\}$ называют множество

$$\lim A_n = \limsup A_n = \liminf A_n,$$

при условии, разумеется, что выполнено равенство $\limsup A_n = \liminf A_n$. Если последнее равенство не выполнено, говорят, что последовательность событий $\{A_n\}$ не имеет предела. Можно показать, что

$$\mathrm{P}(\liminf A_n) \leqslant \liminf \mathrm{P}(A_n) \leqslant \limsup \mathrm{P}(A_n) \leqslant \mathrm{P}(\limsup A_n), \qquad (\text{A.3})$$

откуда и вытекает (A.1).

A.2. Понятие случайной величины. Математические ожидания. *Случайной величиной* на вероятностном пространстве $[\Omega, \mathcal{F}, \mathrm{P}]$ называется измеримая функция $\xi\colon \Omega \mapsto \mathbb{R}$. Измеримость функции ξ эквивалента тому, что для любого $x \in \mathbb{R}$ существует вероятность события $\xi < x$. Из этого условия и свойств сигма-алгебры событий вытекает, что имеют вероятность также $\{\omega \in \Omega\colon \xi(\omega) \geqslant x\}$, $\{\omega \in \Omega\colon \xi(\omega) = x\}$, $\{\omega \in \Omega\colon x_1 < \xi(\omega) < x_2\}$ и т. д. Равенство

$$F(x) = \mathrm{P}\{\omega \in \Omega\colon \xi(\omega) < x\}, \qquad x \in \mathbb{R},$$

задаёт функцию $F\colon \mathbb{R} \mapsto \mathbb{R}$, которая называется *функцией распределения* случайной величины ξ. Обычно в выражениях, содержащих случайные величины, опускают зависимость от ω и пишут $F(x) = \mathrm{P}(\xi < x)$. Можно также определить функцию распределения равенством $F(x) = \mathrm{P}(\xi \leqslant x)$.

Теорема A.1.
Если ξ — случайная величина, то $\xi + a$ и $a\xi$ — случайные величины для любого (неслучайного) числа $a \in \mathbb{R}$. Также являются случайными величинами ξ^2, сумма ξ_1 и ξ_2 двух случайных величин, заданных на одном и том же вероятностном пространстве $[\Omega, \mathcal{F}, \mathrm{P}]$, и предел последовательности $\{\xi_n\}$ случайных величин, заданных на одном и том же вероятностном пространстве, если сходимость $\xi_n(\omega) \to \xi(\omega)$ имеет место для (почти) всех $\omega \in \Omega$.

Доказательство. Чтобы показать, что какая-либо функция $\eta\colon \Omega \mapsto \mathbb{R}$ есть случайная величина, нужно доказать, что для всех $x \in \mathbb{R}$ множество $\{\omega\colon \eta(\omega) < x\}$ принадлежит сигма-алгебре событий $\mathcal{F}$ или, другими словами, что существует $\mathrm{P}(\eta < x)$ для любого $x \in \mathbb{R}$. Все доказательства основаны на том, что операции дополнения, конечного или счётного объединения и конечного или счётного пересечения множеств не выводят из $\mathcal{F}$.

Очевидно, неравенство $\xi + a < x$ эквивалентно $\xi < x - a$, следовательно,

$$\{\omega\colon \xi(\omega) + a < x\} = \{\omega\colon \xi(\omega) < x - a\}, \qquad \mathrm{P}(\xi + a < x) = \mathrm{P}(\xi < x - a).$$

Далее,

$$a\xi < x \Longleftrightarrow \begin{cases} \xi < x/a, & a > 0, \\ \xi > x/a, & a < 0; \end{cases} \qquad \mathrm{P}(a\xi < x) = \begin{cases} \mathrm{P}(\xi < x/a), & a > 0, \\ \mathrm{P}(\xi > x/a), & a < 0. \end{cases}$$

При $a = 0$ мы имеем $a\xi(\omega) \equiv 0$, поэтому $\mathrm{P}(0 < x) = 0$ при $x \leqslant 0$ и $\mathrm{P}(0 < x) = 1$ при $x > 0$.

Для ξ^2 рассуждения также очевидны: $\mathrm{P}(\xi^2 < x) = \mathrm{P}(-\sqrt{x} < \xi < \sqrt{x}\,)$ при $x \geqslant 0$ и $\mathrm{P}(\xi^2 < x) = 0$ при $x < 0$.

Прежде чем доказывать утверждение для $\xi_1 + \xi_2$, покажем, что если ξ_1 и ξ_2 — две случайные величины, заданные на одном и том же вероятностном пространстве, то $\xi_1 < \xi_2$ — событие. Действительно, если ω таково, что $\xi_1(\omega) < \xi_2(\omega)$, то найдётся рациональное число r, при котором $\xi_1(\omega) < r < \xi_2(\omega)$. Верно и обратное включение: если $\xi_1(\omega) < r < \xi_2(\omega)$ при некотором r, то $\xi_1(\omega) < \xi_2(\omega)$. Множество рациональных чисел счётно, пронумеруем их произвольным образом как $r_1, r_2, \ldots$. Тогда

$$\mathrm{P}\big(\omega\colon \xi_1(\omega) < \xi_2(\omega)\big) = \mathrm{P}\left(\bigcup_k \{\omega\colon \xi_1(\omega) < r_k < \xi_2(\omega)\}\right),$$

где вероятность в правой части данного равенства существует, поскольку множества

$$\{\omega\colon \xi_1(\omega) < r_k < \xi_2(\omega)\} = \{\omega\colon \xi_1(\omega) < r_k\} \cap \{\omega\colon \xi_2(\omega) > r_k\}$$

принадлежат сигма-алгебре событий $\mathcal{F}$. Отсюда следует, что $\xi_1 + \xi_2$ — случайная величина, ибо $\mathrm{P}(\xi_1 + \xi_2 < x) = \mathrm{P}(\xi_1 < -\xi_2 + x)$ и, как показано выше, $-\xi_2 + x$ есть случайная величина при любом фиксированном $x \in \mathbb{R}$.

Если $\{\xi_n\}$ — последовательность случайных величин и $\xi_n(\omega) \to \xi(\omega)$ для каждого $\omega \in \Omega$, то

$$\{\omega\colon \xi(\omega) < x\} = \bigcup_{k\in\mathbb{N}} \bigcup_{m\in\mathbb{N}} \bigcap_{n>m} \{\omega\colon \xi_n(\omega) < x - 1/k\}. \tag{A.4}$$

В самом деле, если $\xi(\omega) < x$, то найдётся $k \in \mathbb{N}$ такое, что $\xi(\omega) < x - 1/k$. Далее, в силу сходимости $\xi_n(\omega) \to \xi(\omega)$ найдётся такое $m \in \mathbb{N}$, что для всех $n > m$ имеет место неравенство $|\xi_n(\omega) - \xi(\omega)| < x - 1/k - \xi(\omega)$. Тогда $\xi_n(\omega) < x - 1/k$ для всех $n > m$. Данные рассуждения показывают, что если $\xi(\omega) < x$, то ω принадлежит множеству в правой части равенства (A.4). Наоборот, если найдётся $k \in \mathbb{N}$ такое, что $\xi_n(\omega) < x - 1/k$ при всех n, бо́льших некоторого $m \in \mathbb{N}$, то

$$\xi(\omega) = \lim_{n\to\infty} \xi_n(\omega) = \lim_{\substack{n\to\infty,\\ n>m}} \xi_n(\omega) \leqslant x - 1/k < x,$$

откуда следует, что любой элементарный исход ω, принадлежащий множеству в правой части равенства (A.4), принадлежит и множеству в левой части равенства. Таким образом, множества в (A.4) совпадают. Поскольку в силу свойств сигма-алгебры множество в правой части равенства (A.4) является событием, множество в левой части также принадлежит сигма-алгебре событий, т. е. имеет вероятность. Теорема доказана.

Из доказанной теоремы следует, что

$$\sum_{k=1}^{n} \xi_k, \qquad \sum_{k=1}^{\infty} \xi_k = \lim_{n\to\infty} \sum_{k=1}^{n} \xi_k$$

(если ряд сходится для каждого для каждого $\omega \in \Omega$) суть случайные величины.

Математическим ожиданием случайной величины ξ называется число

$$\mathrm{E}\xi = \int_{\mathbb{R}} x\, dF(x), \tag{A.5}$$

где в правой части стоит интеграл Лебега–Стилтьеса. Считается, что математическое ожидание существует, если интеграл сходится абсолютно. Интеграл Лебега–Стилтьеса от функции $g(x) = x$, $x \in [a, b)$, определяется как предел последовательности интегральных сумм вида

$$\sum_{k=1}^{n} x_k^*\, \mu[x_{k-1}, x_k), \qquad \mu[x_{k-1}, x_k) = \mathrm{P}(\xi \in [x_{k-1}, x_k)) = \big(F(x_k) - F(x_{k-1})\big),$$

где $a \leqslant x_0 < x_1 < \dots < x_n < b$ — произвольные точки на полуинтервале $[a, b)$ и x_k^* — произвольные точки на каждом из полуинтервалов $[x_{k-1}, x_k)$. Заметим, что в этих формулах возникает специфическая мера $\mu(\,\cdot\,)$ множества (в данном случае полуинтервала) на действительной оси. Далее нужно потребовать существование предела последовательности интегральных сумм при

$$\Delta_n = \max_{1\leqslant k\leqslant n} (x_k - x_{k-1}) \to 0, \qquad n \to \infty$$

и независимость его от разбиения полуинтервала $[a, b)$ и выбора точек x_k^*. Этот предел и даёт интеграл Лебега-Стильтьеса (A.5)

Интеграл по всей действительной прямой определяется как обычно:

$$\int_{\mathbb{R}} x\,dF(x) = \lim_{\substack{a\to+\infty,\\ b\to-\infty}} \int_a^b x\,dF(x).$$

Поскольку $F(x_k) - F(x_{k-1}) = \mathrm{P}(x_k \leqslant \xi < x_{k-1})$, мы можем записать равенство в бесконечно малых $dF(x) = \mathrm{P}(x \leqslant \xi < x + dx)$.

Чтобы найти математическое ожидание случайной величины $\eta = g(\xi)$, заданной как некоторая функция от ξ, то по определению мы должны написать

$$\mathrm{E}g(\xi) = \int_{\mathbb{R}} y\,dF_\eta(y), \qquad F_\eta(y) = \mathrm{P}(\eta < y).$$

Но можно также заметить, что мера интервала задаётся как

$$\mu[y_{k-1}, y_k) = \mathrm{P}(y_{k-1} \leqslant g(\xi) < y_k) = \mathrm{P}(\xi \in g^{(-1)}[y_{k-1}, y_k)).$$

где $g^{(-1)}[y_{k-1}, y_k) = \{x\colon y_{k-1} \leqslant g(x) < y_k\}$. Таким образом, мы имеем для интегральных сумм

$$\sum_{k=1}^{n} y_k^*\,\mu[y_{k-1}, y_k) = \sum_{k=1}^{n} g(x_k^*)\mathrm{P}(\xi \in g^{(-1)}[y_{k-1}, y_k)),$$

Если теперь перейти к пределу, то мы естественным образом можем записать

$$\mathrm{E}g(\xi) = \int_{\mathbb{R}} g(x)\,dF(x), \tag{A.6}$$

конечно, требуя при этом абсолютную сходимость интеграла. Тем самым мы определим интеграл Лебега–Стильтеса от функции $g(\,\cdot\,)$ достаточно общего вида.

Дисперсией случайной величины ξ называется величина $\mathrm{D}\xi = \mathrm{E}(\xi - \mathrm{E}\xi)^2$.

Говорят, что набор случайных величин $\xi_1, \ldots, \xi_n$ имеет *совместное распределение*, если для любых $x_1, \ldots, x_n \in \mathbb{R}$ существует вероятность $\mathrm{P}(\xi_1 < x_1, \ldots, \xi_n < x_n)$. Функция $F\colon \mathbb{R}^n \mapsto \mathbb{R}$, заданная равенством $F(x_1, \ldots, x_n) = \mathrm{P}(\xi_1 < x_1, \ldots, \xi_n < x_n)$, называется *совместной функцией распределения* случайных величин $\xi_1, \ldots, \xi_n$. Случайные величины $\xi_1, \ldots, \xi_n$ называются *независимыми*, если

$$F(x_1, \ldots, x_n) = \mathrm{P}(\xi_1 < x_1) \ldots \mathrm{P}(\xi_n < x_n) = F_1(x_1) \ldots F_n(x_n),$$

где $F_k(\cdot)$ — функция распределения случайной величины ξ_k, $k = 1, \ldots, n$.

Если наделить множество $\mathbb{R}^n$ структурой n-мерного векторного, точнее евклидова, пространства $\mathcal{R}^n$ и считать, что все случайные величины $\xi_1, \ldots, \xi_n$ заданы на одном и том же множестве Ω элементарных исходов, то $\xi = \langle \xi_1, \ldots, \xi_n \rangle\colon \Omega \mapsto \mathcal{R}^n$

называется n-мерным *случайным вектором.* Математическое ожидание случайного вектора ξ есть вектор $\mathrm{E}\xi = \langle \mathrm{E}\xi_1, \ldots, \mathrm{E}\xi_n \rangle \in \mathcal{R}^n$, координаты которого суть математические ожидания координат случайного вектора ξ. Аналогом дисперсии служит *матрица ковариаций* Σ размера $n \times n$ с элементами

$$\Sigma_{ij} = \mathrm{E}(\xi_i - \mathrm{E}\xi_i)(\mathrm{E}\xi_j - \xi_j). \tag{A.7}$$

Математическое ожидание в правой части равенства (A.7) называется *коэффициентом ковариации* случайных величин ξ_i и ξ_j и обычно обозначается как $\mathrm{cov}(\xi_i, \xi_j)$. Видно, что матрица ковариаций симметрична, $\Sigma_{ij} = \Sigma_{ji}$.

A.3. Свойства математического ожидания. Важными для наших рассуждений свойствами математического ожидания являются следующие.

• Для любых (неслучайных) чисел $b, a_1, \ldots, a_n$ и случайных величин $\xi_1, \ldots, \xi_n$

$$\mathrm{E}(b + a_1\xi_1 + \cdots + a_n\xi_n) = b + a_1\mathrm{E}\xi_1 + \cdots + a_n\mathrm{E}\xi_n, \tag{A.8}$$

если математические ожидания в правой части равенства существуют (линейность математического ожидания).

• Если случайные величины $\xi_1, \ldots, \xi_n$ независимы, то

$$\mathrm{E}(\xi \ldots \xi_n) = \mathrm{E}\xi_1 \ldots \mathrm{E}\xi_n, \tag{A.9}$$

если, опять же, математические ожидания в правой части равенства существуют. Из этих двух свойств легко вывести, что

$$\mathrm{E}\,\mathrm{const} = \mathrm{const}, \quad \mathrm{D}\,\mathrm{const} = 0, \qquad \mathrm{D}\xi = \mathrm{E}\xi^2 - \mathrm{E}^2\xi, \quad \mathrm{cov}(\xi_i, \xi_j) = \mathrm{E}\xi_i\xi_j - \mathrm{E}\xi_i \cdot \mathrm{E}\xi_j.$$

Кроме того, $\mathrm{cov}(\xi_i, \xi_j) = 0$, если случайные величины ξ_i, ξ_j независимы.

• Из свойств интеграла (A.5) вытекает, что $|\mathrm{E}\xi| \leqslant \mathrm{E}|\xi|$.

• Если $\mathrm{P}(\xi \geqslant 0) = 1$, то $\mathrm{E}\xi \geqslant 0$. Отсюда следует, что:
если $M_1 \leqslant \xi(\omega) \leqslant M_2$ для (почти) всех $\omega \in \Omega$, т. е. $\mathrm{P}(M_1 \leqslant \xi \leqslant M_2) = 1$, то $M_1 \leqslant \mathrm{E}\xi \leqslant M_2$;
если $\xi_1(\omega) \leqslant \xi_2(\omega)$ для (почти) всех $\omega \in \Omega$, т. е. $\mathrm{P}(\xi_1 \leqslant \xi_2) = 1$, то $\mathrm{E}\xi_1 \leqslant \mathrm{E}\xi_2$;
для любой случайной величины $\mathrm{D}\xi \geqslant 0$ или, эквивалентно, $\mathrm{E}\xi^2 \geqslant \mathrm{E}^2\xi$ (если, конечно, математические ожидания существуют).

• Пусть $x = \langle x_1, \ldots, x_n \rangle$ и $z = \langle z_1, \ldots, z_n \rangle$ — произвольные неслучайные векторы из $\mathcal{R}^n$. Тогда для математического ожидания $\mathrm{E}\xi = \langle \mathrm{E}\xi_1, \ldots, \mathrm{E}\xi_n \rangle \in \mathcal{R}^n$ и ковариационной матрицы Σ с элементами $\Sigma_{ij} = \mathrm{cov}(\xi_i, \xi_j)$ случайного вектора

$\xi = \langle \xi_1, \ldots, \xi_n \rangle$ в силу линейности математического ожидания имеют место равенства

$$\begin{aligned}
(\mathrm{E}\xi, x) &= \sum_{i=1}^{n} (\mathrm{E}\xi)_i x_i = \sum_{i=1}^{n} \mathrm{E}\xi_i \, x_i = \mathrm{E} \sum_{i=1}^{n} \xi_i x_i = \mathrm{E}(\xi, x), \\
(\Sigma z, x) &= \sum_{i=1}^{n} (\Sigma x)_i z_i = \sum_{i=1}^{n} \sum_{j=1}^{n} \Sigma_{ij} x_j z_i = \sum_{i,j=1}^{n} \mathrm{E}(\xi_i - \mathrm{E}\xi_i)(\mathrm{E}\xi_j - \xi_j) x_j z_i = \\
&= \mathrm{E} \left(\sum_{j=1}^{n} x_j (\xi_j - \mathrm{E}\xi_j) \cdot \sum_{i=1}^{n} z_i (\xi_i - \mathrm{E}\xi_i) \right) = \mathrm{E}(x, \xi - \mathrm{E}\xi)(z, \xi - \mathrm{E}\xi).
\end{aligned} \tag{A.10}$$

Равенства $(\mathrm{E}\xi, x) = \mathrm{E}(\xi, x)$ и $(\Sigma z, x) = \mathrm{E}(x, \xi - \mathrm{E}\xi)(z, \xi - \mathrm{E}\xi)$ в бесконечномерном случае мы примем за определение математического ожидания и ковариационного оператора. Из последнего равенства ещё раз получаем симметричность ковариационной матрицы, $(\Sigma z, x) = (\Sigma x, z)$, а также то, что она неотрицательно определена: подставляя в последнее равенство $z = x$, имеем $(\Sigma x, x) = \mathrm{E}(x, \xi - \mathrm{E}\xi)^2 \geqslant 0$.

Удобным для исследования свойств математического ожидания является аппарат условных распределений. Пусть B — некоторое событие ненулевой вероятности. *Условной вероятностью* события A (точнее, вероятностью события A при условии, что произошло событие B) называется величина

$$\mathrm{P}(A \,|\, B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}.$$

Особенностью условной вероятности является то, что она удовлетворяет всем аксиомам и свойствам обычной вероятности, если в них заменить Ω на B. Если события $B_1, B_2, \ldots$ удовлетворяет следующим условиям:

$$B_k \cap B_{k'} = \varnothing \text{ при } k \neq k', \qquad \mathrm{P}(B_1 \cup B_2 \cup \cdots) = 1, \qquad \mathrm{P}(B_k) \neq 0, \tag{A.11}$$

то справедлива *формула полной вероятности*: для любого события A

$$\mathrm{P}(A) = \sum_k \mathrm{P}(A \,|\, B_k) \mathrm{P}(B_k).$$

Количество событий B_k может быть конечным или счётным, кроме того, условие $\mathrm{P}(B_k) \neq 0$ можно опустить, считая, что если $\mathrm{P}(B_k) = 0$, то соответствующее слагаемое в сумме равно нулю.

Применим это понятие к распределению случайной величины: заменим в определении $F(x) = \mathrm{P}(\xi < x)$ вероятность на условную, получим *условную функцию распределения* $F(x \,|\, B) = \mathrm{P}(\xi < x \,|\, B)$, точнее функцию распределения случайной величины ξ при условии, что произошло событие B. Условное математическое ожидание определяется как интеграл по условному распределению:

$$\mathrm{E}(\xi \,|\, B) = \int_{\mathbb{R}} x\, dF(x \,|\, B),$$

где, очевидно,

$$dF(x \,|\, B) = \mathrm{P}(x \leqslant \xi < x + dx \,|\, B) = \frac{\mathrm{P}\big(\{\omega\colon x \leqslant \xi(\omega) < x + dx\} \cap B\big)}{\mathrm{P}(B)}.$$

Условное математическое ожидание обладает всеми свойствами обычного математического ожидания, в частности теми, которые мы привели выше.

В ряде случаев удобнее умножить условное математическое ожидание на $P(B)$ и рассматривать величину

$$\mathrm{E}(\xi; B) = \mathrm{E}(\xi \,|\, B)\mathrm{P}(B) = \int_{\mathbb{R}} x\, \mathrm{P}\big(\{\omega\colon x \leqslant \xi(\omega) < x + dx\} \cap B\big). \tag{A.12}$$

Очевидно, что если события $B_1, \ldots, B_n$ удовлетворяют условиям (A.11), то мы можем написать аналог формулы полной вероятности:

$$\mathrm{E}\xi = \sum_{k=1}^{n} \mathrm{E}(\xi \,|\, B_k)\mathrm{P}(B_k) = \sum_{k=1}^{n} \mathrm{E}(\xi; B_k). \tag{A.13}$$

С помощью формулы (A.13) легко доказать следующие неравенства:

$$\mathrm{P}(\xi^2 \geqslant \varepsilon^2) \leqslant \frac{\mathrm{E}\xi^2}{\varepsilon^2}, \qquad \mathrm{P}(|\xi - \mathrm{E}\xi| \geqslant \varepsilon) \leqslant \frac{\mathrm{D}\xi}{\varepsilon^2}, \qquad \varepsilon > 0. \tag{A.14}$$

Первое из них выводится на основе тривиальной цепочки соотношений

$$\mathrm{E}\xi^2 = \mathrm{E}(\xi^2 \,|\, \xi^2 \geqslant \varepsilon^2)\mathrm{P}(\xi^2 \geqslant \varepsilon^2) + \mathrm{E}(\xi^2 \,|\, \xi^2 < \varepsilon^2)\mathrm{P}(\xi^2 < \varepsilon^2) \geqslant \varepsilon^2\mathrm{P}(\xi^2 \geqslant \varepsilon^2), \tag{A.15}$$

поскольку $\mathrm{E}(\xi^2 \,|\, \xi^2 \geqslant \varepsilon^2) \geqslant \varepsilon^2$ и $\mathrm{E}(\xi^2 \,|\, \xi^2 < \varepsilon^2) \geqslant 0$. Чтобы обосновать последние неравенства, достаточно воспользоваться для условных математических ожиданий $\mathrm{E}(\xi^2 \,|\, B)$ следующим свойством (аналогичным свойству $\mathrm{E}\xi^2$): если $\xi^2(\omega) \geqslant M$ для всех $\omega \in B$, то $\mathrm{E}(\xi^2 \,|\, B) \geqslant M$. Понятно, что мы применили это свойство при $M = 0$ во втором неравенстве, а при $M = \varepsilon^2$ — в первом, поскольку в этом слагаемом в роли события B выступает именно $\big\{\omega\colon \xi^2(\omega) \geqslant \varepsilon^2\big\}$. Второе неравенство в (A.14) (*неравенство Чебышёва*) получается при замене в первом неравенстве ξ на $\xi - \mathrm{E}\xi$.

Из неравенства Чебышёва следует, что если $\mathrm{D}\xi = 0$, то $\mathrm{P}(\xi = \mathrm{E}\xi) = 1$. В частности, если во втором соотношении в (A.10) $\Sigma x = 0$, то $\mathrm{E}(x, \xi - \mathrm{E}\xi)^2 = 0$, при этом с учётом первой цепочки равенств в (A.10) мы имеем $\mathrm{E}(\xi - \mathrm{E}\xi, x) = (\mathrm{E}\xi, x) - (\mathrm{E}\xi, x) = 0$, следовательно, $(\xi - \mathrm{E}\xi, x) = 0$ с вероятностью единица.

A.4. Сходимость случайных величин. Последовательностью случайных величин назовём упорядоченный счётный набор случайных величин $\xi_1, \xi_2, \ldots$ таких, что для любого натурального n существует совместное распределение случайных величин $\xi_1, \ldots, \xi_n$. Говорят, что последовательность $\{\xi_n\}$ *сходится с вероятностью единица* (*почти наверное*) к случайной величине ξ, если

$$\mathrm{P}\{\omega \in \Omega\colon\ \xi_n(\omega) \to \xi(\omega)\} = 1 \tag{A.16}$$

(здесь мы предполагаем, все члены последовательности и предельная случайная величина заданы на одном и том же Ω). Говорят, что последовательность $\{\xi_n\}$ *сходится по вероятности* к случайной величине ξ, если для любого $\varepsilon > 0$ имеет место сходимость числовой последовательности вероятностей

$$\mathrm{P}\{|\xi_n - \xi| \geqslant \varepsilon\} \underset{n\to\infty}{\longrightarrow} 0. \tag{A.17}$$

Нетрудно показать, что сходимость с вероятностью единица влечёт сходимость по вероятности. Условие $\xi_n(\omega) \to \xi(\omega)$ означает, что для всякого $\varepsilon > 0$ найдётся $N = N(\omega)$ такое, что для всех $n > N$ имеет место неравенство $|\xi_n(\omega) - \xi(\omega)| < \varepsilon$. Это можно записать как

$$\omega \in \bigcap_{\varepsilon>0} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \bigl\{\omega\colon\ |\xi_n(\omega) - \xi(\omega)| < \varepsilon\bigr\} = \bigcap_{\varepsilon>0} \liminf A_n(\varepsilon),$$

где мы ввели краткое обозначение $\bigl\{\omega\colon\ |\xi_n(\omega) - \xi(\omega)| < \varepsilon\bigr\} = A_n(\varepsilon)$ и использовали определение нижнего предела в (A.2). Тогда условие (A.16) принимает вид

$$\mathrm{P}\Bigl(\bigcap_{\varepsilon>0} \liminf A_n(\varepsilon)\Bigr) = 1, \qquad \varepsilon > 0.$$

Понятно, что

$$\mathrm{P}\Bigl(\bigcap_{\varepsilon>0} \liminf A_n(\varepsilon)\Bigr) \leqslant \mathrm{P}\bigl(\liminf A_n(\varepsilon)\bigr)$$

для любого фиксированного $\epsilon > 0$. Также мы имеем цепочку неравенств (A.3):

$$\mathrm{P}\bigl(\liminf A_n(\varepsilon)\bigr) \leqslant \liminf \mathrm{P}\bigl(A_n(\varepsilon)\bigr) \leqslant \limsup \mathrm{P}\bigl(A_n(\varepsilon)\bigr) \leqslant \mathrm{P}\bigl(\limsup A_n(\varepsilon)\bigr).$$

Объединяя три последних соотношения и учитывая $\mathrm{P}\bigl(\limsup A_n(\varepsilon)\bigr) \leqslant 1$, мы видим, что

$$1 \leqslant \liminf \mathrm{P}\bigl(A_n(\varepsilon)\bigr) \leqslant \limsup \mathrm{P}\bigl(A_n(\varepsilon)\bigr) \leqslant 1,$$

откуда следует, что $P\big(A_n(\varepsilon)\big) = \mathrm{P}(|\xi_n(\omega) - \xi(\omega)| < \varepsilon) \to 1$ для любого $\varepsilon > 0$, и эта сходимость, очевидно, эквивалентна (A.17).

Сходимость случайных величин, вообще говоря, не влечёт сходимость их математических ожиданий, однако при некоторых дополнительных предположениях сходимость математических ожиданий есть следствие сходимости по вероятности. Докажем теорему, являющуюся одним из вариантов знаменитой теоремы Лебега о мажорируемой сходимости.

Теорема A.2.
Пусть последовательность случайных величин $\{\xi_n\}$ сходится к случайной величине ξ по вероятности. Пусть существует случайная величина η такая, что $|\xi_n| \leqslant \eta$ для всех $n = 1, 2, \ldots$ с вероятностью единица. Если существуют математические ожидания $\mathrm{E}\eta$, $\mathrm{E}\xi$ и $\mathrm{E}\xi_n$, $n = 1, 2, \ldots$, то $\mathrm{E}|\xi_n - \xi| \to 0$ при $n \to \infty$.

Доказательство. Сделаем несколько предварительных замечаний общего характера. Условие $B_1 \subset B_2$ влечёт $\mathrm{P}(A \cap B_1) \leqslant \mathrm{P}(A \cap B_2)$ для любого события A, в частности

$$\mathrm{P}\big(\{\omega\colon x \leqslant \xi(\omega) < x + dx\} \cap B_1\big) \leqslant \mathrm{P}\big(\{\omega\colon x \leqslant \xi(\omega) < x + dx\} \cap B_2\big). \tag{A.18}$$

Пусть случайная величина α неотрицательна с вероятностью единица, тогда из соотношений (A.18) получаем

$$\begin{aligned}\mathrm{E}(\alpha; B_1) = \int_{\mathbb{R}_+} x\, \mathrm{P}\big(\{\omega\colon x \leqslant \xi(\omega) < x + dx\} \cap B_1\big) \leqslant \\ \leqslant \int_{\mathbb{R}_+} x\, \mathrm{P}\big(\{\omega\colon x \leqslant \xi(\omega) < x + dx\} \cap B_2\big) = \mathrm{E}(\xi; B_2).\end{aligned} \tag{A.19}$$

С другой стороны, если $\alpha \leqslant \beta$ с вероятностью единица, то аналогично свойству обычного математического ожидания мы имеем $\mathrm{E}(\alpha \,|\, B) \leqslant \mathrm{E}(\beta \,|\, B)$ и

$$\mathrm{E}(\alpha; B) = \mathrm{E}(\alpha \,|\, B) P(B) \leqslant \mathrm{E}(\beta \,|\, B) P(B) \leqslant \mathrm{E}(\alpha; B) \tag{A.20}$$

для любого события B.

Теперь воспользуемся полученными оценками для доказательства теоремы. Применим формулу (A.13) к случайной величине $\alpha = |\xi_n - \xi|$: для любых $\Delta > 0$ и $M > \Delta$ справедливо равенство

$$\begin{aligned}\mathrm{E}|\xi_n - \xi| = \mathrm{E}(|\xi_n - \xi|; 0 \leqslant |\xi_n - \xi| < \Delta) + \\ + \mathrm{E}(|\xi_n - \xi|; \Delta \leqslant |\xi_n - \xi| < M) + \mathrm{E}(|\xi_n - \xi|; |\xi_n - \xi| > M).\end{aligned} \tag{A.21}$$

Рассмотрим каждое из слагаемых по отдельности. Имеем для любого n

$$\mathrm{E}_1(n,\Delta) = \mathrm{E}(|\xi_n - \xi|; 0 \leqslant |\xi_n - \xi| < \Delta) \leqslant \Delta. \tag{A.22}$$

Далее, из условий $|\xi_n| \leqslant \eta$ и $\xi_n \to \xi$, справедливых с вероятностью единица, т. е. для почти всех $\omega \in \Omega$, следует, что $|\xi| \leqslant \eta$ и $|\xi_n - \xi| \leqslant 2\eta$ с вероятностью единица. Тогда условие $|\xi_n - \xi| > M$ влечёт $2\eta > M$, другими словами, мы имеем включение $\{\omega\colon |\xi_n(\omega) - \xi(\omega)| > M\} \subset \{\omega\colon \eta > M/2\}$. Отсюда для любого n в силу соотношений (A.19) и (A.20)

$$\begin{aligned}\mathrm{E}_3(n,M) &= \mathrm{E}(|\xi_n - \xi|; |\xi_n - \xi| > M) \leqslant \mathrm{E}(|\xi_n - \xi|; \eta > M/2) \leqslant \mathrm{E}(2\eta; \eta > M/2) = \\ &= 2\int_{\mathbb{R}_+} y\,\mathrm{P}(y \leqslant \eta \leqslant y + dy; \eta > M) = \\ &= 2\int_M^{+\infty} y\,\mathrm{P}(y \leqslant \eta \leqslant y + dy) = 2\int_M^{+\infty} y\,dF_\eta(y) \underset{M\to\infty}{\longrightarrow} 0.\end{aligned} \tag{A.23}$$

Здесь введено обозначение $F_\eta(\cdot)$ для функции распределения случайной величины η и использовано условие существования $\mathrm{E}\eta$, т. е. сходимость интеграла $\int_{\mathbb{R}_+} y\,dF_\eta(y)$. Заметим, что вследствие $|\xi_n - \xi| \leqslant 2\eta$ сходимость $\mathrm{E}_3(n,M) \to 0$ при $M \to +\infty$ равномерная по n, другими словами, величина $\mathrm{E}_3(n,M)$ может быть сделана сколь угодно малой, если взять $M > M_0$, где M_0 не зависит от n.

Наконец, аналогично (A.23)

$$\begin{aligned}\mathrm{E}_2(n,\Delta,M) &= \mathrm{E}(|\xi_n - \xi|; \Delta \leqslant |\xi_n - \xi| < M) = \int_\Delta^M z\,dF_{|\xi_n - \xi|}(z) \leqslant \\ &\leqslant M\int_\Delta^M dF_{|\xi_n - \xi|}(z) \leqslant M\int_\Delta^\infty dF_{|\xi_n - \xi|}(z) = M\mathrm{P}(|\xi_n - \xi| > \Delta),\end{aligned} \tag{A.24}$$

где $F_{|\xi_n - \xi|}(\cdot)$ — функция распределения случайной величины $|\xi_n - \xi|$.

Теперь мы готовы доказать утверждение теоремы. Пусть $\varepsilon > 0$ произвольно. Возьмём $\Delta = \varepsilon/3$, тогда вследствие (A.22)

$$\mathrm{E}_1(n,\Delta) < \frac{\varepsilon}{3} \quad \text{при всех } n = 1, 2, \ldots\,.$$

В силу (A.23) найдётся номер M, который определяется только значением ε, такой что

$$\mathrm{E}_3(n,M) < \frac{\varepsilon}{3} \quad \text{при всех } n = 1, 2, \ldots\,.$$

По условию $\xi_n \to \xi$ по вероятности, т. е. $P(|\xi_n - \xi| > \Delta) \to 0$ при $n \to \infty$. Для данного значения $\Delta = \varepsilon/3$ найдётся номер N, который определяется величинами

ε, $\Delta = \varepsilon/3$ и $M = M(\varepsilon)$, т. е. в конечном итоге — величиной ε, такой что $P(|\xi_n - \xi| > \Delta) < \varepsilon/3M$ при всех $n > N$, тогда с учётом (A.24)

$$\mathrm{E}_2(n, \Delta, M) < \frac{\varepsilon}{3} \quad \text{при } \ n > N, \quad N = N(\varepsilon).$$

Объединяя полученные оценки, получаем

$$\mathrm{E}|\xi_n - \xi| = \mathrm{E}_1(n, \Delta) + \mathrm{E}_2(n, \Delta, M) + \mathrm{E}_3(n, M) < \varepsilon \quad \text{при } \ n > N(\varepsilon).$$

Теорема доказана.

Понятно, что утверждение теоремы будет тем более верным, если мы потребуем вместо сходимости $\xi_n \to \xi$ по вероятности сходимость с вероятностью единица. Кроме того, $|\mathrm{E}\xi_n - \mathrm{E}\xi| = |\mathrm{E}(\xi_n - \xi)| \leqslant \mathrm{E}|\xi_n - \xi|$, поэтому при выполнении условий теоремы мы также имеем сходимость $\mathrm{E}\xi_n \to \mathrm{E}\xi$.

ПРИЛОЖЕНИЕ B
НЕКОТОРЫЕ ТЕОРЕМЫ ФУНКЦИОНАЛЬНОГО АНАЛИЗА

В этом приложении мы приведём некоторые хорошо известные теоремы функционального анализа, в основном касающиеся свойств линейных операторов, действующих в бесконечномерных действительных сепарабельных гильбертовых пространствах.

Определим классы операторов, рассматривающиеся в настоящем курсе, и приведём их базовые свойства.

• Множество $\mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ *замкнутых плотно определённых* линейных операторов, действующих из пространства $\mathcal{R}$ в пространство $\widetilde{\mathcal{R}}$.

У оператора $A \in \mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ область определения всюду плотна, $\overline{\mathcal{D}(A)} = \mathcal{R}$, и график $\Gamma(A)$,

$$\Gamma(A) = \big\{\langle x, y\rangle \in \mathcal{R} \otimes \widetilde{\mathcal{R}}\colon\ x \in \mathcal{D}(A),\ y = Ax\big\},$$

есть замкнутое линейное многообразие. Другими словами,

$$\begin{cases} \{x_n\} \subset \mathcal{D}(A), \\ x_n \to x, \\ Ax_n \to y \end{cases} \qquad \Longrightarrow \qquad \begin{cases} x \in \mathcal{D}(A), \\ y = Ax. \end{cases}$$

Если A замкнут и плотно определён, то у него существует сопряжённый оператор A^*, также являющийся замкнутым и плотно определённым. При этом

$$\Gamma^{\perp}(A) = \big\{\langle x, y\rangle \in \mathcal{R} \otimes \widetilde{\mathcal{R}}\colon\ y \in \mathcal{D}(A^*),\ x = -A^*y\big\}.$$

Справедливо операторное равенство[5] $A^{**} = A$. Если линейный оператор A замкнут, то его нуль-пространство тоже замкнуто. Заметим, что в общем случае у замкнутого оператора ни $\mathcal{D}(A)$, ни $\mathcal{R}(A)$ не являются замкнутыми линейными многообразиями.

• Банахово пространство $\mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ *всюду определённых ограниченных* линейных операторов, действующих из пространства $\mathcal{R}$ в пространство $\widetilde{\mathcal{R}}$.

[5] Равенство $A = B$ для операторов означает, что выполнены условия $\mathcal{D}(A) = \mathcal{D}(B)$ и $Ax = Bx$ для всех $x \in \mathcal{D}(A) = \mathcal{D}(B)$.

У оператора $A \in \mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ область определения совпадает со всем пространством, $\overline{\mathcal{D}(A)} = \mathcal{R}$, и найдётся некоторая постоянная $M \in \mathbb{R}$ такая, что выполнено условие $\|Ax\| \leqslant M\|x\|$ для всякого $x \in \mathcal{R}$. На множестве операторов, удовлетворяющих данным условиям, вводятся естественные линейные операции

$$(\alpha A + \beta B)x \stackrel{\text{def}}{=} \alpha Ax + \beta Bx, \qquad \alpha, \beta \in \mathbb{R}, \tag{B.1}$$

и норма

$$\|A\| = \sup_{\|x\| \leqslant 1} \|Ax\| = \sup_{\|x\|=1} \|Ax\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}. \tag{B.2}$$

Нормированное линейное пространство $\mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ является полным.

Условие $\|A\| < \infty$ ограниченности оператора A эквивалентно условию его непрерывности: $A(\lim x_n) = \lim Ax_n$ или, другими словами,

$$\begin{cases} \{x_n\} \subset \mathcal{D}(A), \\ x_n \to x, \\ x \in \mathcal{D}(A) \end{cases} \qquad \Longrightarrow \qquad \begin{cases} Ax_n \to y, \\ y = Ax \end{cases} .$$

Для любого $x \in \mathcal{R}$ имеет место оценка $\|Ax\| \leqslant \|A\| \cdot \|x\|$.

Ограниченный (непрерывный) оператор переводит:
ограниченную последовательность в ограниченную, а именно, если $\|x_n\| \leqslant M_1$, то $\|Ax_n\| \leqslant M_2$;
сходящуюся последовательность в сходящуюся: если $x_n \to x$, то $Ax_n \to Ax$;
фундаментальную последовательность в фундаментальную: если $\|x_n - x_m\| \to 0$ при $n, m \to \infty$, то $\|Ax_n - Ax_m\| \to 0$.

Любой оператор A, ограниченный на незамкнутом линейном многообразии $\mathcal{D}(A)$, можно доопределить по непрерывности на $\overline{\mathcal{D}(A)}$ без увеличения нормы: для любого $x = \lim x_n \in \overline{\mathcal{D}(A)}$, где $\{x_n\} \subset \mathcal{D}(A)$, мы определяем действие оператора $\bar{A}$ равенством

$$\bar{A}x = \bar{A}(\lim x_n) \stackrel{\text{def}}{=} \lim Ax_n.$$

При этом

$$\|\bar{A}\| = \sup_{\substack{\|x\| \leqslant 1, \\ x \in \overline{\mathcal{D}(A)}}} \|\bar{A}x\| = \sup_{\substack{\|x\| \leqslant 1, \\ x \in \mathcal{D}(A)}} \|Ax\| = \|A\|.$$

Оператор $\bar{A}$ есть замыкание оператора A, т. е. $\Gamma(\bar{A}) = \overline{\Gamma(A)}$.

Если $A \in \mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, то существует $A^* \in \mathbb{B}(\widetilde{\mathcal{R}} \mapsto \mathcal{R})$.

Условия замкнутости и ограниченности оператора в общем случае не влекут одно другое, однако замкнутый оператор ограничен тогда и только тогда, когда его область определения замкнута (теорема о замкнутом графике), в частности при $\mathcal{D}(A) = \mathcal{R}$.

• Банахово пространство $\mathbb{CO}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ *всюду определённых компактных* линейных операторов, действующих из пространства $\mathcal{R}$ в пространство $\widetilde{\mathcal{R}}$.

Оператор $A \in \mathbb{L}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ называется компактным, если он любую ограниченную последовательность $\{x_n\}$ элементов из своей области определения переводит в компактную последовательность. Компактность означает, что в последовательности $\{Ax_n\}$ найдётся фундаментальная подпоследовательность $\{Ax_{n_k}\}$. Компактные операторы также называют *вполне непрерывными*.

Если оператор компактен, то он ограничен. В случае $\dim \mathcal{R} = \infty$ обратное утверждение неверно. Множество $\mathbb{CO}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, если его наделить линейными операциями (B.1) и нормой (B.2), образует замкнутое линейное многообразие (линейное подпространство) в $\mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$.

Если A — компактный оператор и B — ограниченный оператор, то AB и BA также являются компактными операторами. Оператор $A \in \mathbb{CO}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ не может иметь ограниченного обратного, если $\dim \mathcal{R} = \infty$.

Оператор A компактен тогда и только тогда, когда он всякую слабо фундаментальную последовательность переводит в последовательность, фундаментальную по норме:

$$\begin{cases} \{x_n\} \subset \mathcal{D}(A), \\ \big|(x_n, z) - (x_m, z)\big| \underset{n,m\to\infty}{\longrightarrow} 0, \quad z \in \mathcal{R}, \end{cases} \qquad \Longrightarrow \qquad \|Ax_n - Ax_m\|\big| \underset{n,m\to\infty}{\longrightarrow} 0,$$

для $A \in \mathbb{CO}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ это эквивалентно тому, что если $(x_n, z) \to (x, z)$ для всякого $z \in \mathcal{R}$, то $\|Ax_n - Ax\| \to 0$.

Если $A \in \mathbb{CO}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, то существует $A^* \in \mathbb{CO}(\widetilde{\mathcal{R}} \mapsto \mathcal{R})$.

• Гильбертово пространство $\mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ *всюду определённых* линейных операторов *Гильберта–Шмидта*, действующих из пространства $\mathcal{R}$ в пространство $\widetilde{\mathcal{R}}$.

Всюду определённый оператор A называется оператором Гильберта–Шмидта, если $\sum_k \|Ae_k\|^2 < \infty$ для ОНБ $\{e_k\}$ пространства $\mathcal{R}$. Значение суммы не зависит от выбора ОНБ. На множестве всюду определённых операторов Гильберта–Шмидта можно ввести линейные операции (B.1), а также скалярное произведение и норму

$$(A, B)_2 \overset{\text{def}}{=} \sum_k (Ae_k, Be_k), \qquad \|A\|_2 \overset{\text{def}}{=} \sqrt{\sum_k \|Ae_k\|^2}, \tag{B.3}$$

численные значения которых не зависят о выбора ОНБ; при этом пространство $\mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ оказывается полным по норме $\|\cdot\|_2$. Справедливо неравенство, связывающее норму (B.2) (которую часто называют спектральной) и нормой (B.3) Гильберта–Шмидта: $\|A\| \leqslant \|A\|_2$.

Если $A \in \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, то существует $A^* \in \mathbb{H}(\widetilde{\mathcal{R}} \mapsto \mathcal{R})$.

Формулы (B.3) могут быть записаны с использованием понятия следа оператора:

$$(A, B)_2 = \mathrm{Tr}(AB^*) = \mathrm{Tr}(A^*B) = \mathrm{Tr}(BA^*) = \mathrm{Tr}(B^*A),$$
$$\|A\|_2^2 = \mathrm{Tr}(AA^*) = \mathrm{Tr}(AA^*).$$

Всякий оператор Гильберта–Шмидта является компактным. В случае $\dim \mathcal{R} = \infty$ обратное утверждение неверно.

Если рассматривать введённые пространства операторов просто как множества операторов, то можно написать следующую цепочку включений:

$$\mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}}) \supset \mathbb{B}(\mathcal{R} \mapsto \widetilde{\mathcal{R}}) \supset \mathbb{CO}(\mathcal{R} \mapsto \widetilde{\mathcal{R}}) \supset \mathbb{H}(\mathcal{R} \mapsto \widetilde{\mathcal{R}}),$$

причём в случае бесконечномерного пространства $\mathcal{R}$ все включения являются строгими, а в случае $\dim \mathcal{R} < \infty$ множества операторов в этой цепочке совпадают друг с другом и вообще со множеством всех линейных операторов.

Далее мы докажем несколько хорошо известных утверждений общей теории гильбертовых пространств, а затем ряд теорем, фиксирующих необходимые нам свойства линейных операторов, действующих в бесконечномерных гильбертовых пространствах.

B.1. Теорема Бэра. Множество $\mathcal{M}$ элементов гильбертова пространства $\mathcal{R}$ называется *нигде не плотным*, если оно не плотно ни в каком шаре ненулевого радиуса, т. е. $\overline{\mathcal{M}}$ не содержит в себе шар $\mathcal{K}(x_0, r) = \big\{x \in \mathcal{R}\colon \|x - x_0\| \leqslant r\big\}$ при любом $x_0 \in \mathcal{R}$ и любом $r > 0$ (шар может быть и открытым, другими словами, неравенство в определении шара можно заменить на $\|x - x_0\| < r_0$).

Теорема B.1.
Гильбертово пространство $\mathcal{R}$ не может быть представлено в виде счётного объединения нигде не плотных множеств.

Доказательство. Пусть утверждение неверно, и в гильбертовом пространстве найдется счётный набор множеств $\mathcal{M}_1, \mathcal{M}_2, \ldots$ таких, что возможно представление

$$\mathcal{R} = \bigcup_{n=1}^{\infty} \mathcal{M}_n, \tag{B.4}$$

причём каждое из замкнутых множеств $\overline{\mathcal{M}_n}$ не содержит никакого шара (открытого или замкнутого) ненулевого радиуса.

Рассмотрим произвольный замкнутый шар $\overline{\mathcal{K}}_0 = \overline{\mathcal{K}}(x_0, r_0)$ ненулевого радиуса $r_0 < 1$ и соответствующий открытый шар $\mathcal{K}_0 = \mathcal{K}(x_0, r_0)$. По условию мы имеем $\overline{\mathcal{M}_1} \not\supset \mathcal{K}_0$, поэтому множество $\mathcal{K}_0 \cap (\mathcal{R} \setminus \overline{\mathcal{M}_1})$ непусто и открыто как пересечение двух открытых множеств. Это означает, что найдутся элемент $x_1 \in \mathcal{K}_0 \cap (\mathcal{R} \setminus \overline{\mathcal{M}_1})$ и некоторая его окрестность $\mathcal{K}(x_1, \varepsilon_1)$, $\varepsilon_1 > 0$, такие, что $\mathcal{K}(x_1, \varepsilon_1)$ содержится и в шаре $\mathcal{K}_0$, и в множестве $\mathcal{R} \setminus \overline{\mathcal{M}_1}$. Выберем произвольное положительное число $r_1 < \min(\varepsilon_1, 1/2)$ и положим $\overline{\mathcal{K}}_1 = \overline{\mathcal{K}}(x_1, r_1)$. Тогда

$$\overline{\mathcal{K}}_1 \subset \mathcal{K}(x_1, \varepsilon_1) \subset \mathcal{K}_0 \cap (\mathcal{R} \setminus \overline{\mathcal{M}_1}\,) \subset \overline{\mathcal{K}}_0.$$

Заметим, что по построению шар $\overline{\mathcal{K}}_1$ не имеет общих элементов с множеством $\overline{\mathcal{M}_1}$ и, тем более, общих элементов с множеством $\mathcal{M}_1$. Итак, мы нашли замкнутый шар $\overline{\mathcal{K}}_1 = \overline{\mathcal{K}}(x_1, r_1)$, удовлетворяющий следующим условиям:

$$\overline{\mathcal{K}}_1 \subset \overline{\mathcal{K}}_0, \qquad \overline{\mathcal{K}}_1 \cap \mathcal{M}_1 = \varnothing, \qquad r_1 < \frac{1}{2}. \tag{B.5}$$

Множество $\mathcal{M}_2$ не плотно в шаре $\mathcal{K}_1$, другими словами, $\overline{\mathcal{M}_2} \not\supset \mathcal{K}_1$. По аналогии с предыдущими рассуждениями найдем центр x_2 окрестности $\mathcal{K}(x_2, \varepsilon_2)$, содержащейся в шаре $\mathcal{K}_1$ и в множестве $\mathcal{R} \setminus \overline{\mathcal{M}_2}$, а затем возьмём $0 < r_2 < \min(\varepsilon_2, 1/4)$. Тогда шар $\overline{\mathcal{K}}_2 = \overline{\mathcal{K}}_1(x_2, r_2)$ удовлетворяет условиям

$$\overline{\mathcal{K}}_2 \subset \overline{\mathcal{K}}_1, \qquad \overline{\mathcal{K}}_2 \cap \mathcal{M}_2 = \varnothing, \qquad r_2 < \frac{1}{4}. \tag{B.6}$$

Рассматривая далее множества $\mathcal{M}_3, \mathcal{M}_4, \ldots$, получим последовательность замкнутых вложенных шаров $\{\overline{\mathcal{K}}_n\}$ таких, что

$$\overline{\mathcal{K}}_{n+1} \subset \overline{\mathcal{K}}_n, \qquad \overline{\mathcal{K}}_n \cap \mathcal{M}_n = \varnothing, \qquad r_n < \frac{1}{2^n} \to 0. \tag{B.7}$$

Эта последовательность удовлетворяет теореме о вложенных шарах, следовательно, найдется элемент $x \in \mathcal{R}$ такой, что $x \in \overline{\mathcal{K}}_n$ для всех $n = 1, 2, \ldots$. По построению для каждого $n = 1, 2, \ldots$ пересечение $\overline{\mathcal{K}}_n$ и множества $\mathcal{M}_n$ пусто, поэтому x не принадлежит ни одному из $\mathcal{M}_n$, что противоречит условию (B.4). Теорема доказана.

B.2. Тождественность слабой и сильной ограниченности последовательности. Будем говорить, что последовательность $\{x_n\}$ элементов гильбертова пространства $\mathcal{R}$ *слабо ограничена*, если для любого элемента $z \in \mathcal{R}$ ограничена числовая последовательность $\{(x_n, z)\}$. Другими словами, найдется константа $C(z)$ такая, что $|(x_n, z)| \leqslant C(z)$ для любого $n = 1, 2, \ldots$. Заметим, что если

$\|x_n\| \leqslant C$ для любого $n = 1, 2, \ldots$, то в силу неравенства Коши–Буняковского $|(x_n, z)| \leqslant C\|z\| = C(z)$, т. е. ограниченность последовательности по норме влечёт слабую ограниченность. Обратное следствие неочевидно, однако тоже имеет место.

Теорема B.2.
Если последовательность $\{x_n\}$ элементов гильбертова пространства $\mathcal{R}$ слабо ограничена, то она ограничена по норме.

Доказательство. Пусть $|(x_n, z)| \leqslant C(z)$ для любого $n = 1, 2, \ldots$. Для каждых натуральных n и k зададим множества

$$\mathcal{M}_{nk} = \big\{z \in \mathcal{R}\colon\ |(x_n, z)| \leqslant k\big\},$$
$$\mathcal{M}_k = \bigcap_{n=1}^{\infty} \mathcal{M}_{nk} = \big\{z \in \mathcal{R}\colon\ |(x_n, z)| \leqslant k \text{ для всех } n = 1, 2, \ldots\big\}.$$

Поскольку последовательность $\{x_n\}$ слабо ограничена, для любого $z \in \mathcal{R}$ найдётся натуральное k такое, что $|(x_n, z)| \leqslant C(z) \leqslant k$ для любого $n = 1, 2, \ldots$. Это означает, что $\mathcal{R} = \bigcup_{k=1}^{\infty} \mathcal{M}_k$.

Легко заметить, что если последовательность $\{z_m\}$ содержится в $\mathcal{M}_k$ (это означает, что $|(x_n, z_m)| \leqslant k$ для всех m и n) и $z_m \to z$ при $m \to \infty$, то

$$|(x_n, z)| = |(x_n, \lim_{m\to\infty} z_m)| = \lim_{m\to\infty} |(x_n, z_m)| \leqslant k, \quad \text{для всех } n = 1, 2, \ldots,$$

другими словами, множество $\mathcal{M}_k$ замкнуто.

В силу того что $\mathcal{R} = \bigcup_{k=1}^{\infty} \mathcal{M}_k$, хотя бы одно из множеств $\mathcal{M}_k$ плотно в некотором замкнутом шаре $\overline{\mathcal{K}}(z_0, r_0)$ по теореме Бэра. Это означает, что

$$\mathcal{M}_{k_0} = \overline{\mathcal{M}_{k_0}} \supset \overline{\mathcal{K}}(z_0, r_0)$$

при некотором k_0.

Другими словами, если $\|z - z_0\| \leqslant r_0$, то $|(x_n, z)| \leqslant k_0$ для всех $n = 1, 2, \ldots$, и последовательность $\{(x_n, z)\}$ ограничена равномерно по $z \in \overline{\mathcal{K}}(x_0, r_0)$. Это означает, что константа k_0, ограничивающая $|(x_n, z)|$, не зависит от z. Отсюда следует, что эта последовательность также равномерно ограничена по $z \in \overline{\mathcal{K}}(0, 1)$. В самом деле, любому элементу z такому, что $\|z\| \leqslant 1$, может быть сопоставлен элемент $z' = r_0 z + z_0$, принадлежащий шару $\overline{\mathcal{K}}(x_0, r_0)$. Тогда можно записать $z = (z' - z_0)/r_0$,

$$|(x_n, z)| = \frac{|(x_n, z') - (x_n, z_0)|}{r_0} \leqslant \frac{|(x_n, z')| + |(x_n, z_0)|}{r_0} \leqslant \frac{2k_0}{r_0} = C$$

для всех $n = 1, 2, \ldots$ и любого z такого, что $\|z\| \leqslant 1$. Заметим, что константа C определяется только параметрами шара $\overline{\mathcal{K}}(z_0, r_0)$ и номером k_0 множества, которые «возникли» из теоремы Бэра. Отсюда

$$\|x_n\| = \frac{(x_n, x_n)}{\|x_n\|} = (x_n, z)\big|_{z=x_n/\|x_n\|} \leqslant C, \qquad n = 1, 2, \ldots .$$

Теорема доказана.

В.3. Слабая полунепрерывность нормы снизу. Будем говорить, что последовательность $\{x_n\}$ элементов гильбертова пространства $\mathcal{R}$ *слабо сходится* к элементу $x \in \mathcal{R}$ и писать $x_n \xrightarrow{\text{w}} x$, если для любого $z \in \mathcal{R}$ числовая последовательность $\{(x_n, z)\}$ сходится к (x, z) при $n \to \infty$. Заметим, что если $x_n \to x$, т. е. $\|x_n - x\| \to 0$, то $|(x_n, z) - (x, z)| \leqslant \|x_n - x\| \cdot \|z\| \to 0$ для любого фиксированного $z \in \mathcal{R}$. Это означает, что сходимость последовательности по норме влечёт слабую сходимость. Обратное следствие неверно. Пример, показывающий это, привести несложно. Пусть $\{e_n\}$ — ОНБ пространства $\mathcal{R}$. Тогда для любого $z \in \mathcal{R}$ последовательность $\{(e_n, z)\}$ стремится к нулю, поскольку должно быть выполнено необходимое условие сходимости ряда $\sum_{n=1}^{\infty}(e_n, z)^2 = \|z\|^2$. Но при этом $\|e_n - e_{n'}\|^2 = \|e_n\|^2 + \|e_{n'}\|^2 = 2$ при $n \neq n'$, и последовательность $\{e_n\}$ не является фундаментальной, поэтому не может иметь предела.

Теорема В.3.
Пусть последовательность $\{x_n\}$ элементов гильбертова пространства $\mathcal{R}$ слабо сходится к элементу $x \in \mathcal{R}$. Тогда

$$\|x\|^2 \leqslant \varliminf_{n\to\infty} \|x_n\|^2. \tag{В.8}$$

Доказательство. Положим

$$\beta^2 = \varliminf_{n\to\infty} \|x_n\|^2.$$

Заметим, что последовательность $\{x_n\}$ слабо сходится (числовая последовательность $\{(x_n, z)\}$ имеет предел при любом z), следовательно, $\{x_n\}$ слабо ограничена (числовая последовательность $\{(x_n, z)\}$ ограничена при любом z). Согласно теореме В.2 это означает, что последовательность $\{x_n\}$ ограничена по норме, $0 \leqslant \|x_n\| \leqslant C$, поэтому нижний предел β^2 конечен.

Рассмотрим подпоследовательность $\{x_{n_k}\}_{k=\overline{1,\infty}}$, для которой имеет место сходимость к нижнему пределу: $\|x_{n_k}\|^2 \to \beta^2$ при $k \to \infty$. Тогда

$$0 \leqslant \|x_{n_k} - x\|^2 = \|x_{n_k}\|^2 - 2(x, x_{n_k}) + \|x\|^2 \xrightarrow[k\to\infty]{} \beta^2 - 2(x, x) + \|x\|^2 = \beta^2 - \|x\|^2,$$

откуда вытекает неравенство $\beta^2 - \|x\|^2 \geqslant 0$, эквивалентное (B.8).

Замечание 1. Если $\|x_n - x\| \to 0$, то $\|x_n\| \to \|x\|$, и можно сказать, что любая норма непрерывна относительно сходимости по этой норме. Неравенство (B.8) свидетельствует о том, что непрерывность нормы относительно слабой сходимости, вообще говоря, не имеет места. Пример, показывающий этот факт, привести очень легко: как мы знаем, если $\{e_n\}$ — ОНБ, то последовательность $\{e_n\}$ слабо сходится к нулевому элементу, но при этом $\|e_n\| = 1$ для всех n; в результате мы имеем $0 = \|\lim e_n\|^2 < \lim \|e_n\|^2 = 1$, где в левой части неравенства предел понимается в слабом смысле. Неравенство (B.8) называют свойством *слабой полунепрерывности нормы снизу.*

B.4. Слабая компактность ограниченной последовательности. Будем говорить, что последовательность $\{x_n\}$ элементов гильбертова пространства $\mathcal{R}$ *слабо компактна*, если для любого $z \in \mathcal{R}$ из числовой последовательности $\{(x_n, z)\}$ можно выделить фундаментальную (или, что то же самое, сходящуюся) подпоследовательность.

Понятно, что если из последовательности $\{x_n\}$ можно выделить сходящуюся по норме подпоследовательность, то эта подпоследовательность будет сходиться и в слабом смысле. Таким образом, компактность влечёт слабую компактность, но обратное неверно (см. рассуждения про ОНБ в предыдущем разделе). Кроме того, из курса математического анализа известно, что из каждой ограниченной числовой последовательности можно выделить сходящуюся подпоследовательность. Если заменить числовую последовательность на последовательность $\{x_n\}$ в произвольном гильбертовом пространстве, то это утверждение также потеряет силу. Контрпример даёт всё тот же ОНБ $\{e_n\}$: одновременно мы имеем $\|e_n\| = 1$, следовательно, последовательность $\{e_n\}$ ограничена, и $\|e_n - e_{n'}\|^2 = \|e_n\|^2 + \|e_{n'}\|^2 = 2$ при $n \neq n'$, поэтому никакая подпоследовательность в $\{e_n\}$ не может быть фундаментальной.

Однако если заменить компактность на слабую компактность, мы получим утверждение, аналогичное теореме математического анализа.

Теорема B.4.
Пусть последовательность $\{x_n\}$ элементов гильбертова пространства $\mathcal{R}$ ограничена, $x_n \leqslant C$ для всех $n = 1, 2, \ldots$. Тогда эта последовательность слабо компактна.

Доказательство. Прежде всего отметим, что «для красоты» можно заменить условие ограниченности на слабую ограниченность.

Пусть $\{e_n\}$ — произвольный фиксированный ОНБ в пространстве $\mathcal{R}$.

Рассмотрим числовую последовательность $\{(x_n, e_1)\}$. В силу ограниченности последовательности $\{x_n\}$ имеем $|(x_n, e_1)| \leqslant \|x_n\| \cdot \|e_1\| \leqslant C$ для всех $n = 1, 2, \ldots$, поэтому из числовой последовательности $\{(x_n, e_1)\}$ можно выделить фундаментальную (сходящуюся) подпоследовательность: $(x_{n_k}, e_1) \to \alpha_1$ при $k \to \infty$, где α_1 — некоторое действительное число. Переобозначим члены подпоследовательности, положив $x_k^{(1)} = x_{n_k}$ для всех $k = 1, 2, \ldots$. Тогда

$$\|x_n^{(1)}\| \leqslant C \text{ для всех } n = 1, 2, \ldots, \qquad (x_n^{(1)}, e_1) \to \alpha_1. \tag{B.9}$$

Числовая последовательность $\{(x_n^{(1)}, e_2)\}$ ограничена, следовательно, содержит сходящуюся подпоследовательность: $(x_{n_k}^{(1)}, e_2) \to \alpha_2$ при $k \to \infty$. Вновь переобозначим члены подпоследовательности и положим $x_k^{(2)} = x_{n_k}^{(1)}$ для всех $k = 1, 2, \ldots$. Тогда имеем

$$\{x_n^{(2)}\} \subset \{x_n^{(1)}\}, \quad \|x_n^{(2)}\| \leqslant C, \qquad (x_n^{(2)}, e_2) \to \alpha_2, \quad (x_n^{(2)}, e_1) \to \alpha_1, \tag{B.10}$$

где последняя сходимость имеет место в силу того, что $\{x_n^{(2)}\}$ — подпоследовательность в $\{x_n^{(1)}\}$.

Действуя аналогично для каждого элемента ОНБ, найдем семейство подпоследовательностей, удовлетворяющее условиям

$$\begin{gathered} \{x_n^{(1)}\} \supset \{x_n^{(2)}\} \supset \cdots \supset \{x_n^{(p)}\} \supset \cdots, \\ \|x_n^{(p)}\| \leqslant C \text{ для всех } n = 1, 2, \ldots \quad \text{и при любом } p = 1, 2 \ldots, \\ (x_n^{(p)}, e_p) \underset{n\to\infty}{\longrightarrow} \alpha_p, \qquad (x_n^{(p)}, e_{p'}) \underset{n\to\infty}{\longrightarrow} \alpha_{p'} \text{ при любом } p' < p. \end{gathered} \tag{B.11}$$

Выделим теперь так называемую диагональную подпоследовательность $\{x_p^{(p)}\}$ и для краткости записи положим $z_p = x_p^{(p)}$ для всех $p = 1, 2, \ldots$. Покажем, что для этой последовательности имеет место сходимость $(z_p, e_m) \to \alpha_m$ при $p \to \infty$ для любого фиксированного $m = 1, 2, \ldots$. В самом деле, все элементы $z_m, z_{m+1}, \ldots$ с номерами, не меньшими m, содержатся в последовательности $\{x_n^{(m)}\}_{n=\overline{1,\infty}}$, потому что $z_p = x_p^{(p)}$ и $\{x_n^{(p)}\}_{n=\overline{1,\infty}} \supset \{x_n^{(m)}\}_{n=\overline{1,\infty}}$ при $p \geqslant m$. Таким образом, $\{z_p\}$ — подпоследовательность в $\{x_n^{(m)}\}_{n=\overline{1,\infty}}$, следовательно, $(z_p, e_m) \to \alpha_m$ при $p \to \infty$.

Итак, мы нашли подпоследовательность $\{z_n\}$ в последовательности $\{x_n\}$ такую, что последовательность $\{(z_n, e_m)\}_{n=\overline{1,\infty}}$ является фундаментальной при любом $m = 1, 2, \ldots$. Покажем, что отсюда следует фундаментальность последовательности $\{(z_n, x)\}$ при любом фиксированном $x \in \mathcal{R}$. Разложим элемент x по

ОНБ, $x = \sum_{m=1}^{\infty}(x, e_m)e_m$, и запишем цепочку соотношений

$$\begin{aligned}\big|(z_n, x) - (z_{n'}, x)\big| &= \left|\left(z_n - z_{n'}, \sum_{m=1}^{\infty}(x, e_m)e_m\right)\right| \leqslant \\ &\leqslant \left|\left(z_n - z_{n'}, \sum_{m=1}^{M}(x, e_m)e_m\right)\right| + \left|\left(z_n - z_{n'}, \sum_{m=M+1}^{\infty}(x, e_m)e_m\right)\right|. \qquad \text{(B.12)}\end{aligned}$$

Рассмотрим второе слагаемое в правой части неравенства (B.12) с учётом вытекающего из ограниченности последовательности $\{z_n\} \subset \{x_n\}$ неравенства треугольника $\|z_n - z_{n'}\| \leqslant \|z_n\| + \|z_{n'}\| \leqslant 2C$:

$$\begin{aligned}\mathrm{II}_M &= \left|\left(z_n - z_{n'}, \sum_{m=M+1}^{\infty}(x, e_m)e_m\right)\right| \leqslant \|z_n - z_{n'}\| \cdot \left\|\sum_{m=M+1}^{\infty}(x, e_m)e_m\right\| = \\ &= \|z_n - z_{n'}\|\left(\sum_{m=M+1}^{\infty}(x, e_m)^2\right)^{1/2} \leqslant 2C\left(\|x\|^2 - \sum_{m=1}^{M}(x, e_m)^2\right)^{1/2}.\end{aligned}$$

Поскольку $\|x\|^2 = \sum_{m=1}^{\infty}(x, e_m)^2$, мы имеем $\mathrm{II}_M \to 0$ при $M \to \infty$. Следовательно, для любого $\varepsilon > 0$ найдется номер $M = M(\varepsilon)$ такой, что $\mathrm{II}_M < \varepsilon/2$.

Рассмотрим теперь первое слагаемое в в правой части (B.12):

$$\mathrm{I}_{n,n'} = \left|\left(z_n - z_{n'}, \sum_{m=1}^{M}(x, e_m)e_m\right)\right| = \left|\sum_{m=1}^{M}(z_n - z_{n'}, e_m)(x, e_m)\right|.$$

Последовательность $\{(z_n, e_m)\}_{n=\overline{1,\infty}}$ фундаментальна при любом m, таким образом, в $\mathrm{I}_{n,n'}$ представлена конечная линейная комбинация фундаментальных последовательностей с фиксированными коэффициентами (x, e_m), $m = 1, \ldots, M$. Поэтому $\mathrm{I}_{n,n'} \to 0$ при $n, n' \to \infty$. Точнее, для любого $\varepsilon > 0$ и любого $M < \infty$ найдется $N = N(\varepsilon, M)$ такой, что $\mathrm{I}_{n,n'}\varepsilon/2$ при $n, n' > N$.

Объединяя два рассмотренных слагаемых, получаем, что для любого $\varepsilon > 0$ найдется номер $M = M(\varepsilon)$, а для него — номер $N = N(\varepsilon, M(\varepsilon)) = N(\varepsilon)$ такой, что при $n, n' > N$

$$\big|(z_n, x) - (z_{n'}, x)\big| \leqslant \mathrm{I}_{n,n'} + \mathrm{II}_M \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Теорема доказана.

Следствие.

В обозначениях доказанной теоремы последовательность $\{z_n\}$ слабо сходится к некоторому элементу $z \in \mathcal{R}$.

Доказательство. Мы имеем $(z_n, e_m) \to \alpha_m$ при $n \to \infty$ для каждого фиксированного $m = 1, 2, \ldots$. Покажем, что $\sum_{m=1}^{\infty}\alpha_m^2 < \infty$, тогда из теоремы Фишера–Рисса вытекает, что существует элемент $z = \sum_{m=1}^{\infty}\alpha_m e_m$.

В силу неравенства Бесселя и ограниченности последовательности $\{z_n\}$ для любого конечного M мы можем записать

$$\sum_{m=1}^{M}(z_n, e_m)^2 \leqslant \|z_n\|^2 \leqslant C, \qquad \sum_{m=1}^{M}\alpha_m^2 = \lim_{n\to\infty}\sum_{m=1}^{M}(z_n, e_m)^2 \leqslant C,$$

и последовательность частичных сумм ряда $\sum_{m=1}^{\infty}\alpha_m^2$ ограничена. Таким образом, ряд сходится, и его сумма не превосходит C. Это означает, что существует элемент $z = \sum_{m=1}^{\infty}\alpha_m e_m$, причём $\|z\|^2 = \sum_{m=1}^{\infty}\alpha_m^2 \leqslant C$.

Зафиксируем произвольный элемент $x \in \mathcal{R}$ и покажем, что $(z_n, x) \to (z, x)$. Имеем аналогично (B.12)

$$\begin{aligned} \big|(z_n, x) - (z, x)\big| &= \left|\left(z_n - z, \sum_{m=1}^{\infty}(x, e_m)e_m\right)\right| \leqslant \\ &\leqslant \left|\left(z_n - z, \sum_{m=1}^{M}(x, e_m)e_m\right)\right| + \left|\left(z_n - z, \sum_{m=M+1}^{\infty}(x, e_m)e_m\right)\right| \leqslant \\ &\leqslant \left|\sum_{m=1}^{M}(x, e_m)(z_n - z, e_m)\right| + \|z_n - z\|\left(\sum_{m=M+1}^{\infty}(x, e_m)^2\right)^{1/2}. \end{aligned} \tag{B.13}$$

В силу неравенства $\|z_n - z\| \leqslant \|z_n\| + \|z\| \leqslant 2C$ и сходимости ряда $\sum_{m=1}^{\infty}(x, e_m)^2$ второе слагаемое в правой части стремится к нулю при $M \to \infty$. Поэтому, выбрав достаточно большое M, мы можем сделать это слагаемое меньше $\varepsilon/2$ для любого $\varepsilon > 0$. Зафиксируем это M и подставим его в первое слагаемое в правой части (B.13). Это слагаемое стремится к нулю при $n \to \infty$ вследствие слабой сходимости последовательности $\{z_n\}$ к z. В результате оно также может быть сделано меньше $\varepsilon/2$ при всех n, больших некоторого $N = N(\varepsilon, M(\varepsilon)) = N(\varepsilon)$. Сходимость $(z_n, x) \to (z, x)$ доказана.

B.5. Слабая замкнутость графика оператора. Замкнутость оператора A означает замкнутость его графика. Докажем более общее свойство — слабую замкнутость произвольного линейного подпространства.

Теорема B.5.
Пусть $\mathcal{L}$ – линейное подпространство в гильбертовом пространстве и последовательность $\{x_n\} \subset \mathcal{L}$ слабо сходится, $x_n \xrightarrow{\mathrm{w}} x$. Тогда $x \in \mathcal{L}$.

Разложим предельный элемент x на ортогональные составляющие: $x = x^\circ + x^\perp$, где $x^\circ \in \mathcal{L}$ и $x^\perp \in \mathcal{L}^\perp$. Тогда

$$(x_n, x) = (x_n, x^\circ + x^\perp) = (x_n, x^\circ) \to (x, x^\circ) = (x^\circ, x^\circ) = \|x^\circ\|^2, \qquad n \to \infty,$$

в силу включения $x_n \in \mathcal{L}$ и слабой сходимости последовательности $\{x_n\}$. С другой стороны,

$$(x_n, x) \to (x, x) = \|x\|^2 = \|x^\circ\|^2 + \|x^\perp\|^2, \qquad n \to \infty.$$

Числовая последовательность $\{x_n\}$ может иметь только один предел, поэтому $\|x^\circ\|^2 + \|x^\perp\|^2 = \|x^\circ\|^2$, следовательно, $x^\perp = 0$ и $x \in \mathcal{L}$. Теорема доказана.

Поскольку у замкнутого оператора график есть линейное подпространство в гильбертовом пространстве $\mathcal{R} \otimes \widetilde{\mathcal{R}}$, получаем необходимую нам теорему.

Теорема B.6.

Если $A \in \mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, то его график есть слабо замкнутое множество.

Заметим, что слабая замкнутость линейного подпространства является частным случаем более общего свойства слабой замкнутости любого выпуклого и замкнутого множества.

B.6. Теорема о связи нуль-пространств и пространств значений для замкнутых операторов.

Теорема B.7.
Если $A \in \mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, то

$$\mathcal{N}(A) = \mathcal{N}(A^*A) = \mathcal{R}^\perp(A^*). \tag{B.14}$$

Доказательство. Докажем первое равенство в (B.14). Пусть $x \in \mathcal{N}(A)$, тогда $Ax = 0$ и $A^*Ax = 0$. Пусть $x \in \mathcal{N}(A^*A)$, тогда

$$0 = (x, A^*Ax) = (A^{**}x, Ax) = (Ax, Ax) = \|Ax\|^2,$$

т. е. $Ax = 0$. Равенство доказано.

Докажем второе равенство в (B.14). Пусть $x \in \mathcal{N}(A)$, тогда для любого элемента $z = A^*y \in \mathcal{R}(A^*)$ имеет место цепочка равенств $(x, z) = (x, A^*y) = (Ax, y) = 0$, т. е. $x \in \mathcal{R}^\perp(A^*)$. Пусть $x \in \mathcal{R}^\perp(A^*)$, тогда $\sup_{y \in \mathcal{D}(A^*)} |(A^*y, x)| = 0 < \infty$, что влечёт включение $x \in \mathcal{D}(A^{**})$. При этом $0 = (A^*y, x) = (y, A^{**}x) = (y, Ax)$ для любого элемента $y \in \mathcal{D}(A^*)$. Поэтому $Ax \in \mathcal{D}^\perp(A^*)$, что из чего следует $Ax = 0$ в силу того, что $\mathcal{D}(A^*)$ плотно в $\widetilde{\mathcal{R}}$. Теорема доказана.

Заменив A на A^* и A^* на $(A^*)^* = A$, получим ещё одну цепочку равенств

$$\mathcal{N}(A^*) = \mathcal{N}(AA^*) = \mathcal{R}^\perp(A). \tag{B.15}$$

Рассматривая ортогональные дополнения к множествам в формулах (B.14), (B.15) и пользуясь тем, что для любого линейного многообразия $\mathcal{L}$ справедлива формула $(\mathcal{L}^{\perp})^{\perp} = \overline{\mathcal{L}}$, получаем из (B.14) и (B.15) следующие формулы:

$$\begin{aligned} \mathcal{N}^{\perp}(A^*A) &= \mathcal{N}^{\perp}(A) = \overline{\mathcal{R}(A^*)}, \\ \mathcal{N}^{\perp}(AA^*) &= \mathcal{N}^{\perp}(A^*) = \overline{\mathcal{R}(A)}. \end{aligned} \tag{B.16}$$

B.7. Теорема фон Неймана.

Теорема B.8.

Пусть $A \in \mathbb{CL}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, *тогда*

1) *оператор* $A^*A + I$ *обратим,*

2) *операторы* $(A^*A + I)^{-1}$, $A(A^*A + I)^{-1}$ *и* $A^*A(A^*A + I)^{-1}$ *определены всюду и ограничены, причём*

$$\begin{aligned} \|(A^*A + I)^{-1}\| &\leqslant 1, \\ \|A(A^*A + I)^{-1}\| &\leqslant \frac{1}{\sqrt{2}}, \\ \|A^*A(A^*A + I)^{-1}\| &\leqslant 1. \end{aligned} \tag{B.17}$$

Доказательство. Докажем обратимость оператора $A^*A + I$. Пусть x есть элемент из $\mathcal{N}(A^*A + I)$, тогда $0 = (x, (A^*A + I)x) = \|x\|^2 + \|Ax\|^2$. Отсюда $x = 0$, другими словами, нуль-пространство оператора (A^*A+I) состоит только из одного нулевого элемента. Таким образом, равенство $(A^*A+I)x_1 = (A^*A+I)x_2$ влечёт $x_1 = x_2$, и оператор $(A^*A + I)$ обратим.

Для доказательства остальных утверждений теоремы воспользуемся замкнутостью $\Gamma(A)$ и разложим $\mathcal{R} \otimes \widetilde{\mathcal{R}}$ в прямую сумму $\Gamma(A) \oplus \Gamma^{\perp}(A)$. С учётом явного вида слагаемых

$$\Gamma(A) = \big\{\langle x, Ax\rangle,\ x \in \mathcal{D}(A)\big\}, \qquad \Gamma^{\perp}(A) = \big\{\langle -A^*y, y\rangle,\ y \in \mathcal{D}(A^*)\big\}$$

и единственности разложения по ортогональным составляющим мы можем утверждать, что для каждого $z \in \mathcal{R}$ найдутся единственный элемент $x_z \in \mathcal{D}(A)$ и единственный элемент $y_z \in \mathcal{D}(A^*)$ такие, что

$$\langle z, 0\rangle = \langle x_z, Ax_z\rangle + \langle -A^*y_z, y_z\rangle, \qquad x_z \in \mathcal{D}(A), \quad y_z \in \mathcal{D}(A^*). \tag{B.18}$$

Другими словами, всюду на $\mathcal{R}$ определены операторы S и Y такие, что $x_z = Sz$ и $y_z = Yz$ для каждого $z \in \mathcal{R}$. Данные операторы, очевидно, линейны и, очевидно, $\mathcal{R}(S) \subset \mathcal{D}(A)$, $\mathcal{R}(Y) \subset \mathcal{D}(A^*)$.

Распишем (B.18) покомпонентно с учётом обозначений $x_z = Sz$ и $y_z = Yz$: имеем систему уравнений

$$\begin{cases} z = Sz - A^*Yz, \\ 0 = ASz + Yz. \end{cases}$$

Поскольку z — произвольный элемент пространства $\mathcal{R}$, второе уравнение этой системы влечёт операторное равенство $-AS = Y$. Поскольку $Yz \in \mathcal{D}(A^*)$, мы можем подставить Yz в первое уравнение и получить равенство

$$z = Sz + A^*ASz = (A^*A + I)Sz$$

или, в силу произвольности z, операторное равенство $I = (A^*A + I)S$. Таким образом, $S = (A^*A + I)^{-1}$. Возвращаясь ко второму уравнению системы, которое, напомним, мы можем записать как $-AS = Y$, имеем $Y = -A(A^*A + I)^{-1}$. Поскольку по построению S, Y — всюду определенные операторы и $\mathcal{R}(Y) \subset \mathcal{D}(A^*)$, операторы $(A^*A+I)^{-1} = S$, $A(A^*A+I)^{-1} = -Y$ и $A^*A(A^*A+I)^{-1} = -A^*Y$ также всюду определены.

Покажем неравенства (B.17). Перепишем соотношение (B.18) с учётом определения $x_z = Sz$ и $y_z = Yz$ операторов S и Y:

$$\langle z, 0\rangle = \langle Sz, ASz\rangle + \langle -A^*Yz, Yz\rangle. \tag{B.19}$$

В силу ортогональности слагаемых имеем

$$\|\langle z, 0\rangle\|^2 = \|\langle Sz, ASz\rangle\|^2 + \|\langle -A^*Yz, Yz\rangle\|^2.$$

Раскрываем нормы ($\|\langle x, y\rangle\|^2 = \|x\|^2 + \|y\|^2$), подставляем в явном виде операторы $S = (A^*A + I)^{-1}$, $Y = -A(A^*A + I)^{-1}$ и переписываем данное равенство как

$$\begin{aligned} \|z\|^2 &= \|Sz\|^2 + \|ASz\|^2 + \|-A^*Yz\|^2 + \|Yz\|^2 = \\ &= \|(A^*A + I)^{-1}z\|^2 + 2\|A(A^*A + I)^{-1}z\|^2 + \|A^*A(A^*A + I)^{-1}z\|^2, \end{aligned}$$

где мы учли, что $\|Yz\| = \|ASz\| = \|A(A^*A + I)^{-1}z\|$. Отсюда для любого $z \in \mathcal{R}$ следуют оценки

$$\begin{aligned} \|z\|^2 &\geqslant \|(A^*A + I)^{-1}z\|^2, \\ \|z\|^2 &\geqslant 2\|A(A^*A + I)^{-1}z\|^2, \\ \|z\|^2 &\geqslant \|A^*A(A^*A + I)^{-1}z\|^2, \end{aligned}$$

из которых, очевидно, вытекают неравенства (B.17). Теорема доказана.

Заменяя A на A^*, получаем утверждение, аналогичное теореме фон Неймана, относительно операторов $(AA^* + \tilde{I}\,)^{-1}$, $A^*(AA^* + \tilde{I}\,)^{-1}$ и $AA^*(AA^* + \tilde{I}\,)^{-1}$.

Если мы заменим оператор A на $\omega^{-1/2}A$, где $\omega > 0$, то получим аналоги теоремы фон Неймана для операторов

$$\begin{gathered}
(\omega^{-1}A^*A+I)^{-1} = \omega(A^*A+\omega I)^{-1},\\
\omega^{-1/2}A(\omega^{-1}A^*A+I)^{-1} = \omega^{1/2}A(A^*A+\omega I)^{-1},\\
\omega^{-1}A^*A(\omega^{-1}A^*A+I)^{-1} = A^*A(A^*A+\omega I)^{-1}
\end{gathered} \tag{B.20}$$

и для операторов, полученных при двух заменах $A \leftrightarrow A^*$ и $I \to \tilde{I}$ в формулах (B.20). Все эти операторы всюду определены, и для них имеют место оценки норм (B.17), которые с учётом (B.20) могут быть переписаны как

$$\begin{aligned}
&\|(A^*A+\omega I)^{-1}\| \leqslant \frac{1}{\omega}, && \|(AA^*+\omega \tilde{I}\,)^{-1}\| \leqslant \frac{1}{\omega},\\
&\|A(A^*A+\omega I)^{-1}\| \leqslant \frac{1}{\sqrt{2\omega^2}}, && \|A^*(AA^*+\omega \tilde{I}\,)^{-1}\| \leqslant \frac{1}{\sqrt{2\omega^2}},\\
&\|A^*A(A^*A+\omega I)^{-1}\| \leqslant 1, && \|AA^*(AA^*+\omega \tilde{I}\,)^{-1}\| \leqslant 1.
\end{aligned} \tag{B.21}$$

B.8. Формула для нормы самосопряжённого оператора.

Теорема B.9.
Если $A \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$ и $A = A^*$, то справедлива следующая формула для вычисления нормы этого оператора:

$$\|A\| = \sup_{\|x\|\leqslant 1} |(Ax,x)|. \tag{B.22}$$

Доказательство. Положим

$$\sup_{\|x\|\leqslant 1} |(Ax,x)| = \tau_A.$$

Легко видеть, что $|(Ax,x)| \leqslant \|Ax\|\cdot\|x\| \leqslant \|A\|\cdot\|x\|\cdot\|x\|$ для любого $x \in \mathcal{R}$, поэтому $\tau_A \leqslant \|A\|$. Нетривиальным является доказательство обратного неравенства. Пусть λ – произвольное отличное от нуля действительное число и x – произвольный ненулевой элемент пространства $\mathcal{R}$. Введем два элемента

$$x_\lambda^+ = \frac{1}{\lambda}x + \lambda Ax, \qquad x_\lambda^- = \frac{1}{\lambda}x - \lambda Ax.$$

Таким образом, $Ax_\lambda^\pm = \lambda^{-1}Ax \pm \lambda A^2x$, кроме того, $(x, A^2x) = (Ax, Ax) = \|Ax\|^2$ в силу самосопряжённости A, отсюда

$$\begin{aligned}
(Ax_\lambda^\pm, x_\lambda^\pm) &= \frac{1}{\lambda^2}(Ax,x) \pm (x, A^2x) \pm (Ax,Ax) + \lambda^2(A^2x, Ax) =\\
&= \frac{1}{\lambda^2}(Ax,x) \pm 2\|Ax\|^2 + \lambda^2(A^2x,Ax),\\
\|x_\lambda^\pm|^2 &= \frac{1}{\lambda^2}\|x\|^2 \pm 2(Ax,x) + \lambda^2\|Ax\|^2.
\end{aligned}$$

Комбинируя каждую из двух пар равенств, получаем

$$(Ax_\lambda^+, x_\lambda^+) - (Ax_\lambda^-, x_\lambda^-) = 4\|Ax\|^2, \qquad \|x_\lambda^+\|^2 + \|x_\lambda^-\|^2 = \frac{2}{\lambda^2}\|x\|^2 + 2\lambda^2\|Ax\|^2. \quad \text{(B.23)}$$

Из определения τ_A вытекает оценка $|(Ax_\lambda^\pm, x_\lambda^\pm)| \leqslant \tau_A \cdot \|x_\lambda^\pm\|^2$, поэтому

$$(Ax_\lambda^+, x_\lambda^+) \leqslant \tau_A \cdot \|x_\lambda^+\|^2, \qquad -(Ax_\lambda^-, x_\lambda^-) \leqslant \tau_A \cdot \|x_\lambda^-\|^2.$$

Складывая эти неравенства и подставляя (B.23), получаем для любого $\lambda \neq 0$

$$4\|Ax\|^2 \leqslant \tau_A \left(\frac{2}{\lambda^2}\|x\|^2 + 2\lambda^2\|Ax\|^2\right).$$

В силу произвольности λ мы можем заменить выражение в круглых скобках в правой части последнего неравенства его максимальным по $\lambda^2 > 0$ значением, которое, как нетрудно видеть, достигается при $\lambda^2 = \frac{\|Ax\|}{\|x\|}$: получаем

$$2\|Ax\|^2 \leqslant \tau_A \cdot \left.\left(\frac{1}{\lambda^2}\|x\|^2 + \lambda^2\|Ax\|^2\right)\right|_{\lambda^2 = \frac{\|Ax\|}{\|x\|}} = 2\tau_A \cdot \|Ax\| \cdot \|x\|.$$

Таким образом, для любого ненулевого элемента $x \in \mathcal{R}$ справедливо неравенство $\|Ax\| \leqslant \tau_A\|x\|$, что означает $\|A\| \leqslant \tau_A$. Объединяя с противоположным неравенством, получаем $\|A\| = \tau_A$.

Правая часть равенства (B.22), очевидно, может быть записана в различных формах:

$$\|A\| = \sup_{\|x\|\leqslant 1} |(Ax, x)| = \sup_{|x\|=1} |(Ax, x)| = \sup_{x\neq 0} \frac{|(Ax, x)|}{\|x\|^2}. \qquad \text{(B.24)}$$

B.9. Квадратный корень из оператора.

На множестве самосопряжённых ограниченных операторов может быть введен так называемый частичный порядок, т. е. отношение «больше–меньше».

Пусть $A, B \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$ — самосопряжённые операторы.

Будем говорить, что A — *неотрицательный* оператор и писать $A \geqslant 0$, если $(Ax, x) \geqslant 0$ для любого $x \in \mathcal{R}$. Будем говорить, что A — *положительный* оператор и писать $A > 0$, если $(Ax, x) > 0$ для любого элемента $x \in \mathcal{R}$, отличного от нулевого. Для самосопряжённых операторов $A, B \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$ отношение $A \geqslant B$ означает, что $A - B \geqslant 0$. Аналогично (с заменой знаков $\geqslant$ и $>$ на соответственно $\leqslant$ и $<$) определяются понятия неотрицательного, положительного операторов и отношения $A \leqslant B$.

В отличие от действительных чисел множество самосопряжённых ограниченных операторов упорядочено частично, т. е. не для любых операторов A, B из этого класса можно записать $A \geqslant B$ или $B \geqslant A$, однако при этом отношение «больше–меньше» для операторов, если оно возможно, обладает привычными свойствами:

- если $A \geqslant B$, то $B \leqslant A$;
- если $A \geqslant 0$, то $\lambda A \geqslant 0$ при $\lambda \geqslant 0$ и $\lambda A \leqslant 0$ при $\lambda \leqslant 0$;
- если $A \geqslant B_1$ и $B \geqslant B_2$, то $A + B \geqslant B_1 + B_2$;
- если $A \geqslant B + C$, то $A - B \geqslant C$;
- если $A \geqslant B \geqslant 0$, то $\|A\| \geqslant \|B\|$.

Доказательство первых четырёх свойств легко получить непосредственно из соответствующих определений. Последнее, пятое, свойство вытекает из формулы (B.24). Заметим, что утверждение, обратное пятому свойству, в общем случае неверно, однако если $\|A\| \leqslant 1$ и $A \geqslant 0$, то $(Ax, x) \leqslant (x, x)$ по той же формуле (B.24), т. е. $A \leqslant I$. Таким образом для неотрицательного самосопряжённого A условие $\|A\| \leqslant 1$ эквивалентно $A \leqslant I$.

Докажем неравенство, обобщающее неравенство Коши–Буняковского.

Пусть $S \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$ — самосопряжённый неотрицательный оператор. Тогда для любых $x, z \in \mathcal{R}$ справедливо неравенство

$$\big|(Sx, z)\big| \leqslant \sqrt{(Sx, x) \cdot (Sz, z)}. \tag{B.25}$$

Совершенно аналогично доказательству неравенства Коши–Буняковского получаем, что при всех действительных λ

$$s(\lambda) = \big(S(x + \lambda z), x + \lambda z\big) = \lambda^2 (Sx, x) + 2\lambda (Sx, z) + (Sz, z) \geqslant 0,$$

таким образом, дискриминант квадратного трёхчлена $s(\cdot)$ неположителен, и мы имеем неравенство $(Sx, z)^2 - (Sx, x) \cdot (Sz, z) \leqslant 0$, которое эквивалентно (B.25).

Замечание 2. Потребуем дополнительно к условиям утверждения строгую положительность оператора S: пусть равенство $(Sx, x) = 0$ справедливо, если и только если $x = 0$. Тогда функционал $\|x\|_S = \sqrt{(Sx, x)}$ удовлетворяет аксиомам нормы. В самом деле, положительная определённость и правило вынесения числового множителя из-под знака нормы очевидны, и для любых $x, z \in \mathcal{R}$

$$\big(S(x + z), x + z\big) \leqslant (Sx, x) + 2\sqrt{(Sx, x)(Sz, z)} + (Sz, z) = \Big(\sqrt{(Sx, x)} + \sqrt{(Sz, z)}\,\Big)^2,$$

откуда в силу неотрицательности левой и правой частей неравенства следует неравенство треугольника $\|x + z\|_S \leqslant \|x\|_S + \|z\|_S$. При $S = I$ неравенство (B.25) переходит в неравенство Коши–Буняковского.

Докажем одну теорему, необходимую нам в дальнейшем и представляющую самостоятельный интерес.

Теорема B.10.

Пусть последовательность $\{A_n\} \subset \mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$ *самосопряжённых операторов удовлетворяет условию*

$$0 \leqslant A_1 \leqslant A_2 \leqslant \cdots \leqslant A_n \leqslant \cdots \leqslant I. \tag{B.26}$$

Тогда существует самосопряжённый оператор $A \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$, $0 \leqslant A \leqslant I$, *такой, что при* $n \to \infty$ *имеет место сходимость* $A_n x \to Ax$ *последовательности* $\{A_n\}$ *к оператору* A *на каждом элементе* $x \in \mathcal{R}$.

Доказательство. Пусть x — произвольный фиксированный элемент из пространства $\mathcal{R}$. Из условия (B.26) следует, что

$$0 \leqslant (A_1 x, x) \leqslant (A_2 x, x) \leqslant \cdots \leqslant (A_n x, x) \leqslant \cdots \leqslant (x, x).$$

Другими словами, числовая последовательность $\{(A_n x, x)\}$ монотонно не убывает и ограничена сверху, поэтому она сходится, следовательно, фундаментальна.

По условиям теоремы оператор $S = A_{n+p} - A_n$ самосопряжён и неотрицателен при всех натуральных n, p, поэтому он удовлетворяет условиям обобщённому неравенству Коши–Буняковского (B.25) при $z = Sx$, следовательно,

$$\|A_{n+p}x - A_n x\|^4 = (Sx, Sx)^2 = (Sx, z)^2 \leqslant (Sx, x) \cdot (Sz, z) = (Sx, x) \cdot (S \cdot Sx, Sx). \tag{B.27}$$

Отметим также, что $\|S\| = \|A_{n+p} - A_n\| \leqslant \|A_{n+p}\| + \|A_n\| \leqslant 2$. Отсюда следует, что первый сомножитель в правой части неравенства (B.27) при $n \to \infty$ бесконечно мал, а второй ограничен:

$$(Sx, x) = \big((A_{n+p} - A_n)x, x\big) \underset{n\to\infty}{\longrightarrow} 0, \qquad (S^2 x, Sx) \leqslant \|S\|^3 \|x\|^2 \leqslant 8\|x\|^2.$$

Таким образом, последовательность $\{A_n x\}$ фундаментальна при каждом фиксированном $x \in \mathcal{R}$, т. е. имеет предел. Соответствие $x \mapsto \lim A_n x$ задаёт оператор A, который, очевидно, линеен в силу свойств предельного перехода. Кроме того, для любых $x, z \in \mathcal{R}$ имеют место равенства

$$(Ax, z) = (\lim_{n\to\infty} A_n x, z) = \lim_{n\to\infty} (A_n x, z) = \lim_{n\to\infty} (x, A_n z) = (x, \lim_{n\to\infty} A_n z) = (x, Az).$$

Это означает, что оператор A самосопряжён. Поскольку для любого x можно записать оценки $0 \leqslant (A_n x, x) \leqslant \|x\|^2$, мы имеем неравенства $0 \leqslant (Ax, x) \leqslant \|x\|^2$, откуда следует, что $0 \leqslant A \leqslant I$. Объединяя воедино все доказанные утверждения, получаем, что оператор A удовлетворяет требованиям теоремы.

Для любого ограниченного самосопряжённого оператора $Y \geqslant 0$ определим множество операторных полиномов с неотрицательными коэффициентами:

$$\mathbb{P}(Y) = \left\{ P = \sum_{k=0}^{m} \alpha_k Y^k,\ m = 0, 1, 2, \dots,\ \ \alpha_k \geqslant 0 \right\}.$$

Заметим, что для любого $j = 1, 2, \dots$ и для любых $x, z \in \mathcal{R}$ можно записать

$$\begin{aligned}
(Y^j x, z) &= (Y^{j-1} x, Yz) = \dots = (x, Y^j z), \\
(Y^{2j} x, x) &= \|Y^j x\|^2 \geqslant 0, \\
(Y^{2j+1} x, x) &= (Y \cdot Y^j x, Y^j x) \geqslant 0,
\end{aligned}$$

т. е. любая натуральная степень неотрицательного самосопряжённого оператора есть самосопряжённый неотрицательный оператор. Отсюда следует самосопряжённость и неотрицательность любого $P \in \mathbb{P}(Y)$. Выпишем это и другие очевидные свойства множества операторных полиномов:

$$\begin{aligned}
&\text{если } P \in \mathbb{P}(Y), \text{ то } P = P^* \geqslant 0; \\
&\text{если } P_1, P_2 \in \mathbb{P}(Y),\ \beta_1, \beta_2 \geqslant 0, \text{ то } \beta_1 P_1 + \beta_2 P_2 \in \mathbb{P}(Y); \\
&\text{если } P_1, P_2 \in \mathbb{P}(Y), \text{ то } P_1 P_2 = P_2 P_1 \in \mathbb{P}(Y).
\end{aligned} \tag{B.28}$$

Перейдем к построению квадратного корня из оператора.

Теорема B.11.
Пусть $A \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$ — самосопряжённый оператор и $0 \leqslant A \leqslant I$. Тогда существует оператор B, удовлетворяющий условиям

$$B^2 = A, \qquad B = B^* \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R}), \qquad 0 \leqslant B \leqslant I. \tag{B.29}$$

Доказательство. Положим $X = I - B$, $Y = I - A$ и перепишем уравнение $B^2 = A$ в виде $(I - X)^2 = I - 2X + X^2 = I - Y$ или, что эквивалентно, $2X - X^2 = Y$.

Перепишем последнее уравнение как $X = (Y + X^2)/2$ и на основе это равенства сформируем итерационную последовательность операторов $\{X_n\}$:

$$X_0 = 0, \qquad X_{n+1} = \frac{Y + X_n^2}{2}, \qquad n = 0, 1, 2, \dots . \tag{B.30}$$

Операторы $X_0, Y \in \mathbb{P}(Y)$, отсюда в силу соотношения (B.30) и свойств (B.28) мы имеем включение $X_n \in \mathbb{P}(Y)$ для любого $n = 0, 1, 2, \dots$. Кроме того,

$$X_n^* = X_n \geqslant 0, \qquad X_{n-1} X_n = X_n X_{n-1} \tag{B.31}$$

и в силу (B.30)

$$X_{n+1} - X_n = \frac{X_n^2 - X_{n-1}^2}{2} = \frac{1}{2}(X_n - X_{n-1})(X_n + X_{n-1}). \tag{B.32}$$

Очевидно, что $X_1 - X_0 = 1/2 \cdot Y \in \mathbb{P}(Y)$. Тогда, применяя с учётом $X_n + X_{n-1} \in \mathbb{P}(Y)$ последовательно формулу (B.32) для $n = 1, 2, \ldots$, получаем, что $X_{n+1} - X_n \in \mathbb{P}(Y)$ для всех $n = 0, 1, 2, \ldots$. Следовательно, $X_{n+1} - X_n \geqslant 0$.

Покажем, что $\|X_n\| \leqslant 1$ и, таким образом, $X_n \leqslant I$:

$$\|X_{n+1}\| \leqslant \frac{\|Y\|}{2} + \frac{\|X_n^2\|}{2} \leqslant \frac{\|Y\|}{2} + \frac{\|X_n\|^2}{2}, \tag{B.33}$$

где мы использовали очевидное соотношение

$$\|X_n^2\| = \|X_n \cdot X_n\| \leqslant \|X_n\| \cdot \|X_n\| = \|X_n\|^2.$$

Из оценки (B.33) вытекает, что $\|X_n\| \leqslant 1$ для всех n, ибо $\|X_0\| = 0$ и $\|Y\| \leqslant 1$. Итак,

$$0 \leqslant X_1 \leqslant X_2 \leqslant \cdots \leqslant X_n \leqslant \cdots \leqslant I.$$

В силу теоремы B.10 найдется оператор $0 \leqslant X = X^* \leqslant I$ такой, что $Xz = \lim X_n z$ для любого $z \in \mathcal{R}$. Отсюда для любого $z \in \mathcal{R}$

$$Xz = \lim_{n\to\infty} X_{n+1} z = \lim_{n\to\infty} \frac{Y + X_n^2}{2} z = \frac{Y + X^2}{2} z,$$

т. е. $X = (Y + X^2)/2$, и оператор $B = I - X$ удовлетворяет условиям (B.29) теоремы. Теорема доказана.

Замечание 3. Условие $A \leqslant I$ не является важным для существования квадратного корня, поскольку для любого $A = A^* \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$ оператор $A_\circ = \|A\|^{-1} \cdot A$ удовлетворяет условиям теоремы. Тогда, если $B_\circ$ — квадратный корень из $A_\circ$, то $B = \sqrt{\|A\|} \cdot B_\circ$ — квадратный корень из A.

ПРИЛОЖЕНИЕ C
ЭЛЕМЕНТЫ СПЕКТРАЛЬНОЙ ТЕОРИИ САМОСОПРЯЖЁННЫХ ОПЕРАТОРОВ

C.1. Основные понятия спектральной теории. Пусть A — ограниченный, всюду определённый, самосопряжённый линейный оператор из $\mathcal{R}$ в $\mathcal{R}$ (размерность пространства $\mathcal{R}$ бесконечна, за исключением случаев, когда иное оговорено особо), а λ — произвольное действительное число. Рассмотрим оператор $A - \lambda I$. Очевидно, это также самосопряжённый оператор из банахова пространства $\mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$. Положим $R_\lambda = (A - \lambda I)^{-1}$ и назовем его резольвентным оператором, или просто *резольвентой.* Существуют три взаимоисключающие возможности: оператор R_λ существует и принадлежит банахову пространству $\mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$, оператор R_λ существует, но не принадлежит банахову пространству $\mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$ и, наконец, оператор R_λ не существует. В точном соответствии с этими тремя случаями введем три класса значений λ:

• действительное число λ называется *регулярной точкой* оператора A, если оператор R_λ существует, всюду определен и ограничен;

• действительное число λ называется *точкой непрерывного спектра* оператора A, если оператор R_λ существует, но не принадлежит пространству $\mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$;

• действительное число λ называется *точкой дискретного спектра* оператора A, если оператор R_λ не существует.

Множество точек непрерывного и дискретного спектра называется *спектром оператора* A. Заметим, что R_λ не существует тогда и только тогда, когда найдется ненулевой элемент $x \in \mathcal{R}$, при котором $Ax = \lambda x$. Видно, что в этом случае мы пришли к хорошо известным понятиям: λ есть *собственное значение*, а элемент x — *собственный вектор* оператора A, отвечающий данному собственному значению λ. Таким образом, при нашей классификации точка дискретного спектра и собственное значение суть эквивалентные понятия[6)].

Приведем несколько очевидных утверждений.

• Собственные векторы самосопряжённого ограниченного оператора, отвечающие различным собственным значениям, ортогональны: это следует из триви-

[6)] В большинстве монографий деление на точки непрерывного и дискретного спектра производится более сложным образом, но для наших целей введённых выше определений достаточно.

альных равенств $(\lambda_1 - \lambda_2)(e_1, e_2) = (Ae_1, e_2) - (Ae_2, e_1) = 0$, справедливых, если оператор A самосопряжён, $Ae_1 = \lambda_1 e_1$ и $Ae_2 = \lambda_2 e_2$. Если $\lambda_1 \neq \lambda_2$, то $(e_1, e_2) = 0$.

• Множество собственных векторов ограниченного самосопряжённого оператора, отвечающих одному собственному значению, образует линейное подпространство, если к этому множеству добавить нулевой элемент. В самом деле, множество $\mathcal{L}_\lambda = \{x\colon Ax = \lambda x\}$, очевидно, линейно; докажем его замкнутость. Если $x_n \to x$ и $x_n \in \mathcal{L}_\lambda$, то в силу непрерывности A имеем $Ax = \lim Ax_n = \lambda \lim x_n = \lambda x$. Подпространство $\mathcal{L}_\lambda$ называется *собственным подпространством*, а его размерность — *кратностью собственного значения*. Кратность собственного значения может быть конечной или бесконечной. Собственное подпространство, отвечающее $\lambda = 0$, совпадает с нуль-пространством оператора A. Всюду далее мы будем нормировать собственные векторы и требовать, чтобы они имели единичную норму.

• Если оператор R_λ существует, но не принадлежит пространству $\mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$ (всюду определенных ограниченных операторов), то имеют место оба «нарушения», а именно $\mathcal{D}(R_\lambda) \neq \mathcal{R}$ и R_λ неограничен. В самом деле, если $R_\lambda = (A - \lambda I)^{-1}$ существует, то $\mathcal{N}(A - \lambda I) = \{0\}$, и в силу формул (B.16), применённых к самосопряжённому оператору $A - \lambda I$,

$$\mathcal{D}^\perp(R_\lambda) = \mathcal{R}^\perp(A - \lambda I) = \mathcal{N}(A - \lambda I) = \{0\},$$

поэтому $\mathcal{D}(R_\lambda)$ плотно в $\mathcal{R}$. С другой стороны, $A - \lambda I \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$ есть замкнутый оператор, следовательно, оператор $R_\lambda = (A - \lambda I)^{-1}$ тоже замкнут, и в силу теоремы о замкнутом графике R_λ ограничен тогда и только тогда, когда $\mathcal{D}(R_\lambda)$ замкнуто. Как мы только что показали, область определения $\mathcal{D}(R_\lambda)$ всюду плотна, таким образом, если $\mathcal{D}(R_\lambda) = \mathcal{R}$, то она замкнута, и оператор R_λ ограничен; если же, напротив, $\mathcal{D}(R_\lambda) \neq \mathcal{R}$, то оператор R_λ не может быть ограниченным. Итак, λ — точка непрерывного спектра, тогда и только тогда, когда оператор R_λ существует и $\mathcal{D}(R_\lambda) = \mathcal{R}(A - \lambda I)$ — (всюду плотное) незамкнутое линейное многообразие в $\mathcal{R}$.

Теорема C.1.

Положим

$$\mu_A(\lambda) = \inf_{\|x\|=1} \|Ax - \lambda x\|. \tag{C.1}$$

Имеют место три взаимоисключающие возможности:

1) $\mu_A(\lambda) > 0$ *тогда и только тогда, когда* λ *— регулярная точка оператора* A;

2) $\mu_A(\lambda) = 0$ *и существует элемент* $e \in \mathcal{R}$, $\|e|| = 1$, *на котором достигается точная нижняя грань, т. е.* $\|Ae - \lambda e\| = \mu_A(\lambda)$, *если и только если* λ *— точка дискретного спектра оператора* A;

3) $\mu_A(\lambda) = 0$ *и точная нижняя грань не достигается, т. е.* $\|Ax - \lambda x\| > \mu_A(\lambda)$ *для всех элементов* $x \in \mathcal{R}$ *единичной нормы, если и только если* λ *— точка непрерывного спектра оператора* A.

Доказательство. Рассмотрим каждое из трёх утверждений.

1. Из определения (C.1) следует, что $\|(A-\lambda I)x\| \geqslant \mu_A(\lambda)\|x\|$ для всех x. Отсюда

$$\|y\| \geqslant \mu_A(\lambda)\|(A-\lambda I)^{-1}y\|, \qquad y = (A-\lambda I)x \in \mathcal{R}(A-\lambda I).$$

Последнее неравенство можно переписать как

$$\|(A-\lambda I)^{-1}y\| \leqslant \frac{1}{\mu_A(\lambda)}\|y\|, \qquad y \in \mathcal{D}(A-\lambda I)^{-1},$$

что эквивалентно $\|R_\lambda\| \leqslant \mu_A^{-1}(\lambda) < \infty$. Таким образом, если $\mu_A(\lambda) > 0$, то оператор R_λ ограничен. Наоборот, если оператор R_λ ограничен, то $\|R_\lambda y\| \leqslant \|R_\lambda\|\|y\|$ для любого $y \in \mathcal{R}(A-\lambda I)$, что эквивалентно $\|(A-\lambda I)x\| \geqslant \|R_\lambda\|^{-1}\|x\|$ для всех x. Следовательно, $\mu_A(\lambda) \geqslant \|R_\lambda\|^{-1} > 0$.

2. Если и только если $0 = \mu_A(\lambda) = \|Ae - \lambda e\|$ при некотором элементе e единичной нормы, то e — собственный вектор и λ — собственное значение.

3. Пусть $\mu_A(\lambda) = 0$, но $\|Ax - \lambda x\| > 0$ для любого элемента $x \in \mathcal{R}$ единичной нормы. Отсюда, очевидно, следует, что $(A-\lambda I)x \neq 0$ для любого $x \neq 0$. Другими словами, в этом случае R_λ существует. Рассмотрим последовательности $\{x_n\} \subset \mathcal{R}$ и $\{y_n\} \subset \mathcal{R}(A-\lambda I)$ элементов единичной нормы, связанные равенствами

$$x_n = \frac{(A-\lambda I)^{-1}y_n}{\|(A-\lambda I)^{-1}y_n\|}, \qquad y_n = \frac{(A-\lambda I)x_n}{\|(A-\lambda I)x_n\|} \tag{C.2}$$

(заметим, что $\mathcal{N}(A-\lambda I)^{-1} = \{0\}$ и $\mathcal{N}(A-\lambda I) = \{0\}$, поэтому знаменатели в (C.2) отличны от нуля). Если некая последовательность $\{x_n\}$ элементов единичной нормы доставляет точную нижнюю грань $\mu_A(\lambda) = 0$ в (C.1), то $0 < \|(A-\lambda I)x_n\| \to 0$. Отсюда $\|y_n\| = 1$, но при этом

$$\|R_\lambda y_n\| = \frac{\|x_n\|}{\|(A-\lambda I)x_n\|} = \frac{1}{\|(A-\lambda I)x_n\|} \underset{n\to\infty}{\longrightarrow} \infty,$$

и оператор R_λ неограничен. Если R_λ существует и неограничен, то найдется такая ограниченная последовательность $\{y_n\} \subset \mathcal{R}(A - \lambda I)$, что $\|R_\lambda y_n\| \to \infty$. Тогда элементы x_n, определенные в (C.2), удовлетворяют условиям

$$(A - \lambda I)x_n = \frac{y_n}{\|R_\lambda y_n\|} \underset{n\to\infty}{\longrightarrow} 0, \qquad \|x_n\| = 1,$$

откуда следует, что $\mu_A(\lambda) = 0$, но точная нижняя грань не достигается в силу обратимости оператора $A - \lambda I$.

На основе доказанного свойства легко получить ещё одно утверждение.

Теорема C.2.
Пусть оператор $A \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$ самосопряжён. Тогда

1) *если $|\lambda| > \|A\|$, то λ — регулярная точка*;

2) *существует точка λ_1 спектра оператора A такая, что $|\lambda_1| = \|A\|$*;

3) *λ_1 есть точка дискретного спектра тогда и только тогда, когда найдется элемент e_1, $\|e_1\| = 1$, при котором $|(Ae_1, e_1)| = \|A\|$.*

Доказательство. Вновь рассмотрим каждое из трёх утверждений.

1. Пусть $|\lambda| > \|A\|$. Запишем тривиальную оценку: для любого $x \in \mathcal{R}$

$$|\lambda| \cdot \|x\| = \|\lambda x\| \leqslant \|\lambda x - Ax\| + \|Ax\| = \|(A - \lambda I)x\| + \|Ax\| \leqslant \|(A - \lambda I)x\| + \|A\| \cdot \|x\|,$$

откуда

$$\|(A - \lambda I)x\| \geqslant \big(|\lambda| - \|A\|\big) \cdot \|x\|. \tag{C.3}$$

Таким образом, в (C.1) мы имеем $\mu_A(\lambda) \geqslant |\lambda| - \|A\| > 0$, и это означает, что λ — регулярная точка.

2. Напомним, что $\|A\| = \sup_{\|x\|=1} |(Ax, x)|$ для любого самосопряжённого оператора $A \in \mathbb{B}(\mathcal{R} \mapsto \mathcal{R})$. Рассмотрим последовательность $\{x_n\}$, удовлетворяющую условиям $\|x_n\| = 1$, $\lim |(Ax_n, x_n)| = \|A\|$. Очевидно, что из числовой последовательности $\{(Ax_n, x_n)\}$ можно выделить подпоследовательность $\{(Ax_{n_k}, x_{n_k})\}$, все члены которой имеют один и тот же знак. При этом данная подпоследовательность сходится к $-\|A\|$ или к $+\|A\|$ в зависимости от этого знака. Обозначим предел этой подпоследовательности через λ_1, $|\lambda_1| = \|A\|$. Имеем, учитывая условие $\|x_{n_k}\| = 1$, равенство $\lambda_1^2 = \|A\|$, оценку $\|Ax_{n_k}\|^2 \leqslant \|A\|^2 \cdot \|x_{n_k}\|^2 = \|A\|^2 = \lambda_1^2$ и сходимость $(Ax_{n_k}, x_{n_k}) \to \lambda_1$,

$$\begin{aligned} \|Ax_{n_k} - \lambda_1 x_{n_k}\|^2 &= \|Ax_{n_k}\|^2 - 2\lambda_1 (Ax_{n_k}, x_{n_k}) + \lambda_1^2 \|x_{n_k}\|^2 \leqslant \\ &\leqslant 2\big(\lambda_1^2 - \lambda_1 (Ax_{n_k}, x_{n_k})\big) \to 2(\lambda_1^2 - \lambda_1^2) = 0, \quad k \to \infty. \end{aligned} \tag{C.4}$$

Отсюда следует, что подпоследовательность $\{x_{n_k}\}$ доставляет точную нижнюю грань $\inf_{\|x\|=1} \|Ax - \lambda_1 x\| = 0$, и λ_1 — точка спектра, $|\lambda_1| = \|A\|$.

3. Пусть λ_1, $|\lambda_1| = \|A\|$, есть точка дискретного спектра (собственное значение), и $Ae_1 = \lambda_1 e_1$, $\|e_1\| = 1$. Тогда $|(Ae_1, e_1)| = |\lambda_1|(e_1, e_1) = \|A\|$. С другой стороны, если существует элемент e_1 единичной нормы, для которого $|(Ae_1, e_1)| = \|A\|$, то

$$\|A\| = |(Ae_1, e_1)| \leqslant \|Ae_1\| \cdot \|e_1\| \leqslant \|A\| \cdot \|e_1\|^2 = \|A\|. \tag{C.5}$$

Отсюда, во-первых, вытекает равенство $|(Ae_1, e_1)| = \|Ae_1\| \cdot \|e_1\|$, что возможно, если и только если при некотором λ_1 выполнено равенство $Ae_1 = \lambda_1 e_1$; во-вторых, мы имеем $\|Ae_1\| = \|A\|$, $\|e_1\| = 1$, и это влечёт $|\lambda_1| = \|A\|$. Теорема доказана.

Из третьего утверждения теоремы следует, что если элемент $e \in \mathcal{R}$ удовлетворяет условиям

$$|(Ae, e)| = \sup_{\|x\|=1} |(Ax, x)|, \quad \|e\| = 1, \tag{C.6}$$

то e — собственный вектор оператора A с собственным значением λ_1, $|\lambda_1| = \|A\|$.

C.2. Спектр самосопряжённого компактного оператора. Рассмотрим самосопряжённый компактный оператор A, разумеется, будем считать его отличным от нулевого. Поскольку всякий компактный оператор ограничен, при рассмотрении будем использовать спектральные свойства, доказанные выше. Посмотрим, какие новые свойства появляются при добавлении условия компактности оператора.

Отметим прежде всего, что $\lambda = 0$ есть точка спектра любого компактного самосопряжённого оператора A, действующего в бесконечномерном пространстве, а именно точка дискретного спектра[7)] в случае вырожденного оператора A и точка непрерывного спектра в случае невырожденного оператора A. В самом деле, если $\mathcal{N}(A)$ содержит ненулевые элементы, то $\lambda = 0$ есть точка дискретного спектра, $Ax = 0 \cdot x$ при некотором $x \neq 0$. Если A невырожден, то он не может иметь ограниченного обратного. Таким образом, оператор $A^{-1} = R_0 = (A - 0 \cdot I)^{-1}$ неограничен, следовательно, $\lambda = 0$ есть точка непрерывного спектра оператора A.

Докажем основное утверждение о структуре спектра компактного оператора.

Теорема C.3.

Все ненулевые точки спектра оператора A являются собственными значениями.

Доказательство. Пусть некоторое действительное число $\lambda \neq 0$ есть точка спектра оператора A. Согласно теореме 1 при этом $\inf_{\|x\|=1} \|Ax - \lambda x\| = 0$. Пусть последовательность $\{x_n\}$ элементов единичной нормы доставляет указанную точную

[7)] Ещё раз напомним о том, что наше определение точки дискретного спектра является сильно упрощённым.

нижнюю грань. Оператор A компактен, последовательность $\{x_n\}$ ограничена, поэтому $\{Ax_n\}$ компактна, т. е. найдется подпоследовательность $\{x_{n_k}\}$, для которой имеют место сходимости

$$d_k = Ax_{n_k} - \lambda x_{n_k} \underset{k\to\infty}{\longrightarrow} 0, \qquad Ax_{n_k} \underset{k\to\infty}{\longrightarrow} y \tag{C.7}$$

(первая вытекает из того, что $\{x_{n_k}\}$ доставляет точную нижнюю грань, вторая соответствует определению компактности A). Поскольку $\lambda \neq 0$, существует

$$\lim_{k\to\infty} x_{n_k} = \lim_{k\to\infty} \frac{d_k - Ax_{n_k}}{\lambda} = \lambda^{-1} y. \tag{C.8}$$

Кроме того, в силу непрерывности A можно утверждать, что $Ax_{n_k} \to \lambda^{-1} Ay$ при $k \to \infty$. Сравнивая со вторым пределом в (C.7), получаем $Ay = \lambda y$. Согласно (C.8) $\lambda^{-1}\|y\| = \lim_{k\to\infty} \|x_{n_k}\| = 1$, поэтому $y \neq 0$, откуда следует, что λ — собственное значение. Теорема доказана.

Итак, спектр самосопряжённого компактного оператора A является чисто дискретным, за исключением, быть может, точки $\lambda = 0$.

Введем понятие, которое будет использовано в дальнейших рассуждениях. Будем говорить, что линейное подпространство $\mathcal{L}$ является *инвариантным* относительно оператора A, если условие $x \in \mathcal{L}$ влечёт $Ax \in \mathcal{L}$. На любом инвариантном подпространстве можно задать оператор $[A]_{\mathcal{L}}$, действующий по правилу $[A]_{\mathcal{L}} x = Ax$ для всех $x \in \mathcal{L}$. Данный оператор называется *сужением оператора* A на подпространство $\mathcal{L}$.

Нетрудно заметить, что $[A]_{\mathcal{L}}$ как линейный оператор из $\mathcal{L}$ в $\mathcal{L}$ обладает теми же свойствами, что и оператор A: если A ограничен, то $[A]_{\mathcal{L}}$ ограничен, причём

$$\|[A]_{\mathcal{L}}\| = \sup_{x\in\mathcal{L},\ \|x\|=1} \|[A]_{\mathcal{L}} x\| = \sup_{x\in\mathcal{L},\ \|x\|=1} \|Ax\| \leqslant \sup_{x\in\mathcal{R},\ \|x\|=1} \|Ax\| = \|A\|; \tag{C.9}$$

если A самосопряжён, то $[A]_{\mathcal{L}}$ также самосопряжён, т. к. для любых $x, z \in \mathcal{L}$

$$\big([A]_{\mathcal{L}} x, z\big) = (Ax, z) = (Az, x) = ([A]_{\mathcal{L}} z, x);$$

если A компактен, то $[A]_{\mathcal{L}}$ также компактен, т. к. если $\{x_n\} \subset \mathcal{L}$ — ограниченная последовательность, то из последовательности $\{[A]_{\mathcal{L}} x_n\} = \{Ax_n\}$ выделяется сходящаяся подпоследовательность, предел которой принадлежит $\mathcal{L}$ в силу замкнутости линейного подпространства.

Примером инвариантного пространства является $\mathcal{L}^{\perp}\{e_1, e_2, \ldots, e_n\}$ при условии, что $e_1, e_2, \ldots, e_n$ — собственные векторы оператора A. В самом деле, в этом случае из $(x, e_k) = 0$ следует $(Ax, e_k) = \lambda_k (x, e_k) = 0$.

Из теоремы 2 вытекает, что вне отрезка $\big[-\|A\|, \|A\|\big]$ точек спектра нет и существует точка спектра λ_1 (с необходимостью являющаяся собственным значением) такая, что $|\lambda_1| = \|A\|$. Пусть e_1 — собственный вектор, отвечающий этому собственному значению. Рассмотрим линейное подпространство $\mathcal{L}_1^\perp = \mathcal{L}^\perp(e_1)$. Как отмечалось выше, сужение $A_1 = [A]_{\mathcal{L}_1^\perp}$ оператора A на инвариантное подпространство $\mathcal{L}_1^\perp$ есть самосопряжённый компактный оператор. Кроме того, согласно (C.9)

$$\|A_1\| \leqslant \|A\| = |\lambda_1|. \tag{C.10}$$

Рассмотрим каждую из двух взаимоисключающих возможностей.

Если $\|A_1\| = 0$, то $A_1 x = Ax = 0$ для любого $x \in \mathcal{L}_1^\perp$, другими словами, $x \in \mathcal{L}_1^\perp$ влечёт $Ax = 0$. Верно и обратное утверждение. Если $x \notin \mathcal{L}_1^\perp$, то $x = (x, e_1)e_1 + x^\perp$, где $x^\perp \in \mathcal{L}_1^\perp$, причём $(x, e_1)e_1 \neq 0$. В силу инвариантности подпространства $\mathcal{L}_1^\perp$ мы можем утверждать, что $Ax^\perp$ ортогонален e_1. Отсюда

$$\|Ax\|^2 = \|\lambda_1(x, e_1)e_1 + Ax^\perp\|^2 = \|\lambda_1(x, e_1)e_1\|^2 + \|Ax^\perp\|^2 \geqslant \lambda_1^2(x, e_1)^2 \neq 0. \tag{C.11}$$

Это означает, что условие $x \notin \mathcal{L}_1$ влечёт $Ax \neq 0$. Таким образом, мы имеем равенство подпространств $\mathcal{L}_1^\perp = \mathcal{L}^\perp(e_1) = \mathcal{N}(A)$.

Если $\|A_1\| \neq 0$, то, применяя к A_1 приведённые выше рассуждения, найдём собственное значение λ_2, $|\lambda_2| = \|A_1\|$. Согласно оценке (C.10) мы имеем $|\lambda_2| \leqslant |\lambda_1|$. Отвечающий λ_2 собственный вектор e_2 по построению имеет единичную норму и принадлежит $\mathcal{L}^\perp(e_1)$. Продолжая аналогичные рассуждения, рассмотрим инвариантное относительно A подпространство $\mathcal{L}_2^\perp = \mathcal{L}^\perp(e_1, e_2)$ и сужение A_2 оператора A на это подпространство. Если норма сужения равна нулю, то $Ax = 0$ для всех $x \in \mathcal{L}_2^\perp$ и, если $x \notin \mathcal{L}_2^\perp$, аналогично (C.11) мы имеем $x = (x, e_1)e_1 + (x, e_2)e_2 + x^\perp$, причём $x \neq x^\perp$, откуда

$$\|Ax\|^2 \geqslant \lambda_1^2(x, e_1)^2 + \lambda_2^2(x, e_2)^2 \neq 0.$$

Отсюда $\mathcal{L}^\perp(e_1, e_2) = \mathcal{N}(A)$. Если $\|A_2\| \neq 0$, то существует собственное значение λ_3, причём по аналогии с (C.10) $|\lambda_3| \leqslant |\lambda_2| \leqslant |\lambda_1|$. Отвечающий λ_3 собственный вектор e_3, очевидно, образует вместе с элементами e_1, e_2 ортонормированную систему: $(e_i, e_k) = \delta_{ik}$ для $i, k = 1, 2, 3$. В этом случае находим инвариантное относительно A подпространство $\mathcal{L}_3^\perp = \mathcal{L}^\perp(e_1, e_2, e_3)$ и повторяем рассуждения заново. Таким образом, получаем следующую альтернативу.

Первая возможность: найдется конечный номер r, при котором собственные векторы $e_1, e_2, \ldots, e_r$ отвечают ненулевым собственным значениям $\lambda_1, \lambda_2, \ldots, \lambda_n$, но при переходе в следующее инвариантное подпространство $\mathcal{L}_r^\perp = \mathcal{L}^\perp(e_1, e_2, \ldots, e_r)$ мы имеем

$$\sup_{x \in \mathcal{L}_r^\perp, \|x\|=1} \|Ax\| = 0. \tag{C.12}$$

Тогда

$$\mathcal{L}^\perp(e_1, e_2, \ldots, e_r) = \mathcal{N}(A), \qquad \overline{\mathcal{R}(A)} = \mathcal{N}^\perp(A) = \mathcal{L}(e_1, e_2, \ldots, e_r), \tag{C.13}$$

другими словами, в этом случае A — оператор конечного ранга r.

Второая возможность: для любого $r = 1, 2, \ldots$ точная верхняя грань в (C.12) положительна. Тогда существует последовательность $\{\lambda_n\}$ ненулевых собственных значений и отвечающая им последовательность $\{e_n\}$ собственных векторов, причём

$$\|A\| = |\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_n| \cdots \tag{C.14}$$

и собственные векторы образуют ортонормированную систему в $\mathcal{R}$. Заметим, что $e_k = \lambda_k^{-1} A e_k \in \mathcal{R}(A)$ для всех $k = 1, 2, \ldots$. В силу взаимной ортогональности e_k в этом случае $\operatorname{rank} A = \dim \mathcal{R}(A) = \infty$. Рассмотрим этот случай подробнее.

Докажем основное свойство последовательности $\{\lambda_n\}$.

Теорема C.4.

Последовательность $\{\lambda_n\}$ не имеет предельных точек, отличных от нуля.

Доказательство. Пусть $\lambda \neq 0$ — предельная точка последовательности $\{\lambda_n\}$, а сходящаяся к λ подпоследовательность есть $\{\lambda_{n_k}\}$. Очевидно, в силу того, что $\lim_{k\to\infty} \lambda_{n_k} \neq 0$, мы можем утверждать, что последовательность $\{\lambda_{n_k}^{-1}\}$ ограничена. Рассмотрим последовательность элементов $z_k = \lambda_{n_k}^{-1} e_{n_k}$, $k = 1, 2, \ldots$. Данная последовательность ограничена, но последовательность элементов $Az_k = e_{n_k}$, $k = 1, 2, \ldots$, некомпактна в силу ортогональности собственных векторов. Противоречие с условием компактности оператора A доказывает утверждение.

Поскольку последовательность $\{|\lambda_n|\}$ монотонно не возрастает, она имеет предел, и в силу предыдущей теоремы очевидно следующее утверждение.

Теорема C.5.

Существует $\lim\limits_{n\to\infty} |\lambda_n| = 0$.

Из основного свойства также имеем такое утверждение.

Теорема C.6.

Все найденные выше собственные значения λ_k имеют конечную кратность.

Доказательство. По построению каждому из λ_k отвечает один собственный вектор, однако собственные значения могут совпадать, т. е. возможно

$$\lambda_{k-1} \neq \lambda_k = \lambda_{k+1} = \ldots = \lambda_{k+p} \neq \lambda_{k+p+1},$$

и в этом случае число p есть кратность собственного значения $\lambda = \lambda_k$. Покажем, что $p < \infty$. Пусть это не так, и существует счётный набор (подпоследовательность) равных друг другу собственных значений, $\lambda_k = \lambda_{k+1} = \cdots$. Но в этом случае мы можем утверждать, что значение $\lambda_k \neq 0$ есть предельная точка последовательности $\{\lambda_n\}$, что противоречит основному свойству.

Докажем равенства, аналогичные (C.13). Отметим сначала, что если у оператора $A \in \mathbb{CO}(\mathcal{R} \mapsto \mathcal{R})$ пространство значений $\mathcal{R}(A)$ бесконечномерно, то оно с необходимостью незамкнуто. В самом деле, сузим пространство $\mathcal{R}$ до $\mathcal{N}^{\perp}(A)$ и зададим оператор $A_\circ$ равенством $A_\circ x = Ax$ для всех $x \in \mathcal{N}^{\perp}(A)$. Очевидно, $A_\circ$ — невырожденный линейный оператор из $\mathcal{N}^{\perp}(A)$ в $\widetilde{\mathcal{R}}$, причём $\mathcal{R}(A_\circ) = \mathcal{R}(A)$. Легко также понять, что $A_\circ$ — компактный оператор, как и оператор A. Поскольку $A_\circ$ компактен, он ограничен и, следовательно, замкнут. Обратный оператор $A_\circ^{-1}$ тоже замкнут. Следовательно, если линейное многообразие $\mathcal{R}(A_\circ) = \mathcal{D}(A_\circ^{-1})$ замкнуто, то по теореме о замкнутом графике оператор $A_\circ^{-1}$ ограничен. Но в бесконечномерном пространстве компактный оператор $A_\circ$ не может иметь ограниченного обратного. Заметим, что приведённые выше рассуждения не используют факта самосопряжённости оператора A, т. е. верны и в случае несамосопряжённого оператора.

Теорема C.7.
Справедливы следующие равенства:

$$\mathcal{N}(A) = \mathcal{L}^{\perp}(e_1, \ldots, e_n, \ldots), \qquad \mathcal{R}(A) = \mathcal{L}(e_1, \ldots, e_n, \ldots). \tag{C.15}$$

Доказательство. Если $Ax = 0$, то $(x, e_j) = \lambda_j^{-1}(Ax, e_j) = 0$ для любого натурального j, т. е. $x \in \mathcal{L}^{\perp}(e_1, \ldots, e_n, \ldots)$. С другой стороны, если ненулевой элемент $z \in \mathcal{L}^{\perp}(e_1, \ldots, e_n, \ldots)$, то для всех $k = 1, 2, \ldots$

$$\frac{\|Az\|}{\|z\|} \leqslant \sup_{\substack{x \in \mathcal{L}^{\perp}(e_1,\ldots,e_k),\\ x \neq 0}} \frac{\|Ax\|}{\|x\|} = \sup_{\substack{x \in \mathcal{L}^{\perp}(e_1,\ldots,e_k),\\ \|x\|=1}} \|Ax\| = |\lambda_{k+1}| \underset{k\to\infty}{\longrightarrow} 0.$$

Отсюда $Az = 0$. Таким образом доказано первое равенство утверждения.

Из доказанного равенства следует, что

$$\overline{\mathcal{R}(A)} = \mathcal{N}^{\perp}(A) = \overline{\mathcal{L}(e_1, \ldots, e_n, \ldots)},$$

что с учётом незамкнутости $\mathcal{R}(A)$ даёт второе равенство в (C.15).

Следующее утверждение показывает, что приведённая выше процедура нахождения собственных значений (как норм последовательных сужений оператора A на ортогональные дополнения к собственным подпространствам) исчерпывает все ненулевые собственные значения как в случае $\operatorname{rank} A = r < \infty$, так и в случае $\operatorname{rank} A = \infty$.

Теорема C.8.
Не существует числа $\lambda \neq 0$, являющегося собственным значением и не совпадающего ни с одним λ_k, найденных выше.

Доказательство. Пусть $Ae = \lambda e$ при некоторых λ и $e \in \mathcal{R}$, $\|e\| = 1$. Из общих свойств собственных векторов следует, что, если $\lambda \neq \lambda_j$, то $(e, e_j) = 0$ при всех $j = 1, 2, \ldots$. Другими словами, $e \in \mathcal{L}^{\perp}(e_1, e_2, \ldots) = \mathcal{N}(A)$, и $\lambda e = Ae = 0$. Отсюда $\lambda = 0$.

Подведем итог. По построению $|\lambda_1| = \|A\|$, а для прочих ненулевых собственных значений величина $|\lambda_k|$ равна норме сужения A на ортогональное дополнение к линейной оболочке $\mathcal{L}(e_1, e_2, \ldots, e_{k-1})$ предыдущих собственных векторов. Собственные значения и собственные векторы удовлетворяют следующим равенствам:

$$\begin{aligned}
|\lambda_1| = |(Ae_1, e_1)| = \|Ae_1\| \quad &= \sup_{x\in\mathcal{R},\ \|x\|=1} |(Ax, x)|,\\
|\lambda_k| = |(Ae_k, e_k)| = \|Ae_k\| \quad &= \sup_{\substack{x\in\mathcal{L}^{\perp}(e_1,\ldots,e_{k-1}),\\ \|x\|=1}} |(Ax, x)| = \sup_{\substack{x\in\mathcal{L}^{\perp}(e_1,\ldots,e_{k-1}),\\ \|x\|=1}} \|Ax\|,
\end{aligned} \tag{C.16}$$

Учитывая теорему B.9 из приложения B мы можем также написать следующие равенства

$$\begin{aligned}
|\lambda_1| = |(Ae_1, e_1)| &= \sup_{x\in\mathcal{R},\ \|x\|=1} \|Ax\| = \|A\|,\\
|\lambda_k| = |(Ae_k, e_k)| &= \sup_{\substack{x\in\mathcal{L}^{\perp}(e_1,\ldots,e_{k-1}),\\ \|x\|=1}} \|Ax\| = \|A_k\|,
\end{aligned} \tag{C.17}$$

где A_k — сужение оператора A на инвариантное подпространство $\mathcal{L}^{\perp}(e_1, \ldots, e_{k-1})$. Если обозначить через $\mathcal{L}^{\perp}(e_1, \ldots, e_r)$, $r \leqslant \infty$, ортогональное дополнение линейной оболочки всех собственных векторов, отвечающих ненулевым собственным значениям, не разделяя случаи конечного или счётного ранга A, то имеет место формула

$$\sup_{\substack{x\in\mathcal{L}^{\perp}(e_1,\ldots,e_r),\\ \|x\|=1}} |(Ax, x)| = \sup_{\substack{x\in\mathcal{L}^{\perp}(e_1,\ldots,e_r),\\ \|x\|=1}} \|Ax\| = 0. \tag{C.18}$$

Отметим ещё одно, чрезвычайно важное для наших рассуждений минимаксное свойство собственных подпространств.

Теорема C.9.
Пусть $A \in \mathbb{CO}(\mathcal{R} \mapsto \mathcal{R})$ — самосопряжённый оператор, причём $A \geqslant 0$. Пусть $\lambda_1, \lambda_2, \ldots$ — его ненулевые собственные значения, $e_1, e_2, \ldots$ — соответствующие им (ортонормированные) собственные векторы, найденные по формулам (C.16). Для любого натурального $m < \operatorname{rank}(A)$ имеет место следующая минимаксная формула:

$$\lambda_{m+1} = \min_{\mathcal{L}:\ \dim \mathcal{L} \leqslant m} \ \sup_{x \in \mathcal{L}^\perp,\ \|x\| \leqslant 1} (Ax, x), \tag{C.19}$$

где минимум вычисляется по всем линейным подпространствам в $\mathcal{R}$, размерность которых не превосходит m.

Доказательство. Пусть $\mathcal{L}_{m+1} = \mathcal{L}(e_1, \ldots, e_{m+1})$ — линейная оболочка первых $m+1$ собственных векторов оператора A. По условию теоремы $m+1$ не превосходит ранга оператора A, поэтому собственные значения $\lambda_1, \ldots, \lambda_{m+1}$ отличны от нуля, более того, в силу условия $A \geqslant 0$, положительны (ибо $\lambda_j = (Ae_j, e_j) \geqslant 0$). Рассмотрим $\mathcal{L}$ — произвольное линейное подпространство размерности $s \leqslant m$ и введем в нём ОНБ $\varphi_1, \ldots, \varphi_s$.

Покажем, что линейное подпространство $\mathcal{L}^\perp \cap \mathcal{L}_{m+1}$ содержит ненулевые элементы. В самом деле условия $z \in \mathcal{L}^\perp = \mathcal{L}^\perp(\varphi_1, \ldots, \varphi_s)$ и $z \in \mathcal{L}_{m+1} = \mathcal{L}(e_1, \ldots, e_{m+1})$ эквивалентны следующим требованиям:

$$(z, \varphi_q) = 0 \ \text{для всех} \ q = 1, \ldots, s, \qquad z = \sum_{j=1}^{m+1} \beta_j e_j.$$

Подставляя второе условие в первое, видим, что мы имеем однородную систему линейных уравнений вида

$$\sum_{j=1}^{m+1} \beta_j (e_j, \varphi_q) = 0, \qquad q = 1, \ldots, s, \tag{C.20}$$

относительно коэффициентов $\beta_1, \ldots, \beta_{m+1}$. Число уравнений этой системы меньше числа неизвестных (по условию $s \leqslant m < m+1$), поэтому система имеет ненулевые решения. Другими словами, найдутся числа $\beta_1, \ldots, \beta_{m+1}$ такие, что $\sum_{j=1}^{m+1} \beta_j^2 \neq 0$, удовлетворяющие (C.20). Поскольку данная система однородна, числа $C\beta_1, \ldots, C\beta_{m+1}$ также являются решениями при любом множителе C. Следовательно, всегда найдётся решение, удовлетворяющее условию $\sum_{j=1}^{m+1} \beta_j^2 = 1$. Это означает, что элемент $z = \sum_{j=1}^{m+1} \beta_j e_j$ принадлежит $\mathcal{L}^\perp \cap \mathcal{L}_{k+1}$ и имеет единичную норму.

Для найденного элемента $z \in \mathcal{L}^\perp$ с $\|z\| = 1$ справедливо очевидное неравенство

$$\sup_{x \in \mathcal{L}^\perp,\ \|x\| \leqslant 1} |(Ax, x)| \geqslant |(Az, z)|. \tag{C.21}$$

С другой стороны, $Ae_j = \lambda_j$, следовательно, $Az = \sum_{j=1}^{m+1} \beta_j \lambda_j e_j$, поэтому в силу ортономированности собственных векторов

$$(Az, z) = \sum_{j=1}^{m+1} \beta_j^2 \lambda_j (e_j, e_j) = \sum_{j=1}^{m+1} \lambda_j \beta_j^2 \geqslant \min_{1 \leqslant j \leqslant m+1} \lambda_j \cdot \sum_{j=1}^{m+1} \beta_j^2 = \lambda_{m+1}, \tag{C.22}$$

где мы воспользовались упорядоченностью $\lambda_1 \geqslant \cdots \geqslant \lambda_{m+1} > 0$ собственных значений и условием нормировки $\|z\|^2 = \sum_{j=1}^{m+1} \beta_j^2 = 1$. Подставляя оценку (C.22) в (C.21), получаем, что для любого линейного подпространства $\mathcal{L}$, $\dim \mathcal{L} = s \leqslant m$,

$$\sup_{x \in \mathcal{L}^\perp,\, \|x\| \leqslant 1} |(Ax, x)| \geqslant \lambda_{m+1}.$$

Для завершения доказательства заметим, что при $\mathcal{L} = \mathcal{L}(e_1, \ldots, e_m)$ в соответствии с (C.16)

$$\lambda_{m+1} = (Ae_{m+1}, e_{m+1}) = \sup_{\substack{x \in \mathcal{L}^\perp(e_1, \ldots, e_m),\\ \|x\| \leqslant 1}} (Ax, x)$$

(мы не пишем в формулах модуль, поскольку $(Ax, x) \geqslant 0$ для любого $x \in \mathcal{R}$). Таким образом, минимум в (C.19) достигается на $\mathcal{L} = \mathcal{L}(e_1, \ldots, e_m)$. Теорема доказана.

C.3. Сингулярные базисы компактных операторов. Откажемся от самосопряжённости оператора. Пусть $A \in \mathbb{CO}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$ — компактный оператор. Положим

$$\begin{aligned} \tilde{r} &= \dim \mathcal{R}(A) \leqslant \infty, \qquad & r &= \dim \mathcal{R}(A^*) \leqslant \infty, \\ n &= \dim \mathcal{N}(A) \leqslant \infty, \qquad & \tilde{n} &= \dim \mathcal{N}(A^*) \leqslant \infty. \end{aligned}$$

Напомним также, что

$$\begin{aligned} \mathcal{N}(A) &= \mathcal{N}(A^*A) = \mathcal{R}^\perp(A^*A) = \mathcal{R}^\perp(A^*), \\ \mathcal{N}(A^*) &= \mathcal{N}(AA^*) = \mathcal{R}^\perp(AA^*) = \mathcal{R}^\perp(A). \end{aligned} \tag{C.23}$$

Оператор A^*A компактен, самосопряжён, неотрицателен (поскольку мы имеем $(A^*Ax, x) = \|Ax\|^2 \geqslant 0$ для любого $x \in \mathcal{R}$), и $\dim \mathcal{R}(A^*A) = \dim \mathcal{R}(A) = r$. Пусть $A^*Ae_k = a_k^2 e_k$, элементы $e_1, \ldots, e_r$ ортонормированы, собственные значения $a_1^2, \ldots, a_r^2$ пронумерованы в порядке убывания, $a_1^2 \geqslant a_2^2 \geqslant \cdots$, положительны в силу $A^*A \geqslant 0$ и удовлетворяют условиям, аналогичным (C.16): при всех $k = 0, 1, 2, \ldots, r$

$$a_{k+1}^2 = \sup_{\substack{x \in \mathcal{L}^\perp(e_1, \ldots, e_k),\\ \|x\| le 1}} (A^*Ax, x) = \sup_{\substack{x \in \mathcal{L}^\perp(e_1, \ldots, e_k),\\ \|x\| le 1}} \|Ax\|^2 \tag{C.24}$$

(для единообразия формул мы положили $\mathcal{R} = \mathcal{L}^{\perp}(e_1, \dots, e_k)$ при $k = 0$ и $a_{r+1}^2 = 0$ в случае $r < \infty$). Для каждого $k = 1, \dots, r$ зададим элементы пространства $\widetilde{\mathcal{R}}$ равенствами

$$\tilde{e}_k = \frac{1}{a_k} A e_k, \text{ где } a_k = +\sqrt{a_k^2}. \tag{C.25}$$

Легко видеть, что $\tilde{e}_k \in \mathcal{R}(A)$ и

$$(\tilde{e}_k, \tilde{e}_j) = a_k/a_j \cdot (e_k, e_j) = \delta_{kj}, \quad A^* \tilde{e}_k = {a_k}^{-1} A^* A e_k = {a_k}^{-1} a_k^2 e_k = a_k e_k$$

при $k = 1, \dots, r$. Кроме того, в силу равенств (C.13) и (C.15)

$$\mathcal{L}^{\perp}(e_1, \dots, e_r) = \mathcal{N}(A^* A) = \mathcal{N}(A), \qquad \overline{\mathcal{L}(e_1, \dots, e_r)} = \overline{\mathcal{R}(A^* A)} = \overline{\mathcal{R}(A)}.$$

Итак, существуют ортонормированные системы $\{e_k\}_{k=\overline{1,r}}$ в $\mathcal{R}$ и $\{\tilde{e}_k\}_{k=\overline{1,r}}$ в $\widetilde{\mathcal{R}}$, для которых

$$A e_k = a_k \tilde{e}_k, \quad A^* \tilde{e}_k = a_k e_k, \qquad k = 1, \dots, r. \tag{C.26}$$

Дополним ортонормированный базис $e_1, \dots, e_r$ линейного многообразия $\mathcal{R}(A^* A)$ произвольными ортонормированными элементами $e_1^0, \dots, e_n^0$, образующими базис подпространства $\mathcal{N}(A) = \mathcal{R}^{\perp}(A^* A)$. В силу (C.23) объединение этих ортонормированных систем образует ОНБ пространства $\mathcal{R}$. Дополним также ортонормированную систему $\tilde{e}_1, \dots, \tilde{e}_r$ из (C.25) произвольными ортонормированными элементами $\tilde{e}_1^0, \tilde{e}_2^0, \dots$ до базиса пространства $\widetilde{\mathcal{R}}$. Покажем, что эти элементы образуют ОНБ подпространства $\mathcal{N}(A^*)$. Из (C.26) следует, что для всех $k = 1, 2, \dots, r$

$$(A^* \tilde{e}_i^0, e_k) = (\tilde{e}_i^0, A e_k) = a_k (\tilde{e}_i^0, \tilde{e}_k) = 0,$$

а для всех $j = 1, 2, \dots, n$ в силу $e_j^0 \in \mathcal{N}(A)$

$$(A^* \tilde{e}_i^0, e_j^0) = (\tilde{e}_i^0, A e_j) = 0.$$

Таким образом, элемент $A^* e_i$ ортогонален всем элементам ОНБ пространства $\mathcal{R}$, и это влечет $A^* \tilde{e}_i^0 = 0$. Отсюда следуте, что $\mathcal{L}(\tilde{e}_1^0, \tilde{e}_2^0, \dots) \subset \mathcal{N}(A^*)$, и поэтому $\overline{\mathcal{L}(\tilde{e}_1^0, \tilde{e}_2^0, \dots)} \subset \mathcal{N}(A^*)$, ибо $\mathcal{N}(A)$ замкнуто.

С другой стороны, $\mathcal{L}(\tilde{e}_1, \dots, \tilde{e}_r) \subset \mathcal{R}(A)$, поэтому

$$\overline{\mathcal{L}(\tilde{e}_1^0, \tilde{e}_2^0, \dots)} = \mathcal{L}^{\perp}(\tilde{e}_1, \dots, \tilde{e}_r) \supset \mathcal{R}^{\perp}(A) = \mathcal{N}(A^*).$$

На основании приведённых включений имеем $\overline{\mathcal{L}(\tilde{e}_1^0, \tilde{e}_2^0, \dots)} = \mathcal{N}(A^*)$, и размерность данного подпространства есть $\tilde{n} \leqslant \infty$.

Заметим, что в силу (C.23)

$$\overline{\mathcal{L}(\tilde{e}_1, \dots, \tilde{e}_r)} = \mathcal{L}^{\perp}(\tilde{e}_1^0, \dots, \tilde{e}_{\tilde{n}}) = \overline{\mathcal{R}(A)}.$$

Отсюда $\tilde{r} = \operatorname{rank} A$ совпадает с $r = \operatorname{rank} A^*$, где в случае оператора бесконечного ранга равенство размерностей многообразий $\mathcal{R}(A)$ и $\mathcal{R}(A^*)$ подразумевает возможность установить между элементами их ортонормированных базисов взаимно однозначное соответствие (см. (C.26)).

Итак, имеем следующий факт. Для любого компактного оператора A ранга $r \leqslant \infty$ найдутся:

- ортонормированный базис $\{e_k\}_{k=\overline{1,r}} \oplus \{e_k^0\}_{k=\overline{1,n}}$ пространства $\mathcal{R}$,
- ортонормированный базис $\{\tilde{e}_k\}_{k=\overline{1,r}} \oplus \{\tilde{e}_i^0\}_{i=\overline{1,\tilde{n}}}$ пространства $\widetilde{\mathcal{R}}$,
- действительные числа $a_1 \geqslant \cdots \geqslant a_r > 0$ такие, что

$$\begin{aligned} Ae_k &= a_k \tilde{e}_k, & A^*\tilde{e}_k &= a_k e_k, \\ Ae_j^0 &= 0, & A^*\tilde{e}_i^0 &= 0 \end{aligned} \tag{C.27}$$

для $k = 1, \ldots, r$ и $j = 1, \ldots, n$, $i = 1, \ldots, \tilde{n}$.

При этом:

- ортонормированные системы $\{e_k\}_{k=\overline{1,r}}$ и $\{\tilde{e}_k\}_{k=\overline{1,r}}$ суть базисы линейных многообразий $\mathcal{R}(A^*)$ и $\mathcal{R}(A)$ соответственно, $r = \dim \mathcal{R}(A) = \dim \mathcal{R}(A^*) = \operatorname{rank} A$;
- ортонормированные системы $\{e_j^0\}_{j=\overline{1,n}}$ и $\{\tilde{e}_i^0\}_{i=\overline{1,\tilde{n}}}$ суть базисы линейных подпространств $\mathcal{N}(A)$ и $\mathcal{N}(A^*)$ соответственно, $n = \dim \mathcal{N}(A)$ и $\tilde{n} = \dim \mathcal{N}(A^*)$;
- размерности базисов удовлетворяют равенствам $r+n = \dim \mathcal{R}$ и $r+\tilde{n} = \dim \widetilde{\mathcal{R}}$ (каждая из упомянутых выше размерностей может быть равна бесконечности).

Базисы в (C.27) называются *сингулярными базисами* оператора A, а числа $a_1, \ldots, a_r$ — *сингулярными числами* оператора A.

Для любого $x \in \mathcal{R}$ можно записать разложение в ряд

$$\begin{aligned} Ax &= A\left(\sum_{k=1}^{r}(x, e_k)e_k + \sum_{j=1}^{n}(x, e_j^0)e_j^0\right) = \\ &= \sum_{k=1}^{r}(x, e_k)Ae_k + \sum_{j=1}^{n}(x, e_j^0)Ae_j^0 = \sum_{k=1}^{r} a_k(x, e_k)\tilde{e}_k + 0. \end{aligned}$$

(в случаях $r = \infty$, $n = \infty$ оператор можно заносить под знак суммы в силу непрерывности A). Ряд

$$Ax = \sum_{k=1}^{r} a_k(x, e_k)\tilde{e}_k \tag{C.28}$$

называется *рядом Шмидта* для оператора $A \in \mathbb{CO}(\mathcal{R} \mapsto \mathcal{R})$. Ряд Шмидта сопряжённого оператора, очевидно, имеет вид

$$A^*y = \sum_{k=1}^{r} a_k(y, \tilde{e}_k)e_k \tag{C.29}$$

для любого элемента $y \in \widetilde{\mathcal{R}}$.

Для всех $k = 0, \ldots, r$ справедливы следующие равенства:

$$a_{k+1} = \sup_{\substack{x \in \mathcal{L}^{\perp}(e_1,\ldots,e_k),\\ \|x\| \leqslant 1}} \|Ax\| = \sup_{\substack{y \in \mathcal{L}^{\perp}(\tilde{e}_1,\ldots,\tilde{e}_k),\\ \|y\| \leqslant 1}} \|A^* y\|. \tag{C.30}$$

В самом деле, первое из равенств в (C.30) есть просто преобразованное равенство (C.24). Докажем второе равенство в (C.30). Пусть $0 \leqslant k < r$ и элемент $y \in \mathcal{L}^{\perp}(\tilde{e}_1, \ldots, \tilde{e}_k)$. Тогда

$$y = \sum_{q=k+1}^{r} (y, \tilde{e}_q)\tilde{e}_q + \sum_{i=1}^{\tilde{n}} (y, \tilde{e}_i^0)\tilde{e}_i^0.$$

Вследствие невозрастания последовательности a_k, $k = 1, 2, \ldots$, с учётом (C.29) можно написать оценку

$$\|A^* y\|^2 = \sum_{q=k+1}^{r} a_k^2 (y, \tilde{e}_k)^2 \leqslant a_{k+1}^2 \|y\|^2, \qquad y \in \mathcal{L}^{\perp}(\tilde{e}_1, \ldots, \tilde{e}_k).$$

Таким образом, вторая точная верхняя грань в (C.30) не превосходит a_{k+1}. С другой стороны, для $y = \tilde{e}_{k+1} \in \mathcal{L}^{\perp}(\tilde{e}_1, \ldots, \tilde{e}_k)$ имеем $\|A^* \tilde{e}_{k+1}\| = a_{k+1}$, т. е. эта точная нижняя грань достигается на элементе $\tilde{e}_{k+1}$.

При $r < \infty$ с учётом $\mathcal{L}^{\perp}(e_1, \ldots, e_r) = \mathcal{N}(A)$ и $\mathcal{L}^{\perp}(\tilde{e}_1, \ldots, \tilde{e}_r) = \mathcal{N}(A^*)$ получаем, что при $k = r$ точные верхние грани в (C.30) равны нулю, и это согласуется с принятым обозначением $a_{r+1} = 0$.

C.4. Минимаксное свойство сингулярных чисел. Пусть $a_1 \geqslant \cdots \geqslant a_r > 0$ — сингулярные числа оператора $A \in \mathbb{CO}(\mathcal{R} \mapsto \widetilde{\mathcal{R}})$, $r \leqslant \infty$, а $\{e_k\}$ и $\{\tilde{e}_k\}$ — части сингулярных базисов, отвечающие этим числам, $Ae_k = a_k \tilde{e}_k$. Из теоремы C.9, заменяя A на A^*A и, соответственно, (Ax, x) на $(A^*Ax, x) = \|Ax^2\|$, получаем следующую теорему.

Теорема C.10.
Для любого натурального $m < \operatorname{rank}(A)$

$$a_{m+1}^2 = \min_{\mathcal{L}\colon\, \dim \mathcal{L} \leqslant n} \; \sup_{x \in \mathcal{L}^{\perp},\, \|x\| \leqslant 1} \|Ax\|^2, \tag{C.31}$$

где минимум вычисляется по всем линейным подпространствам в $\mathcal{R}$, *размерность которых не превосходит* m.

Из этой теоремы вытекает следующее утверждение.

Теорема C.11.

Пусть $\mathbb{V}_m$ — множество всюду определённых линейных операторов из $\mathcal{R}$ в $\widetilde{\mathcal{R}}$, ранг которых не превосходит m. Тогда

$$\min_{V\in\mathbb{V}_m}\|A-V\| = \begin{cases} a_{m+1}, & \text{если } m < \operatorname{rank} A,\\ 0, & \text{если } m \geqslant \operatorname{rank} A. \end{cases} \tag{C.32}$$

Доказательство. Пусть $m < \operatorname{rank} A$ и $V \in \mathbb{V}_m$ — произвольный оператор. Воспользуемся теоремой C.10: поскольку[8] $\dim\mathcal{R}(V^*) = \dim\mathcal{R}(V) = \operatorname{rank} V < m$, для любого оператора $V \in \mathbb{V}_m$

$$\begin{aligned} a_{m+1} &= \min_{\mathcal{L}:\ \dim\mathcal{L}\leqslant m}\ \sup_{x\in\mathcal{L}^\perp,\ \|x\|\leqslant 1}\|Ax\| \leqslant \sup_{x\in\mathcal{R}^\perp(V^*),\ \|x\|\leqslant 1}\|Ax\| = \\ &= \sup_{x\in\mathcal{R}^\perp(V^*),\ \|x\|\leqslant 1}\|Ax-Vx\| \leqslant \sup_{x\in\mathcal{R},\ \|x\|\leqslant 1}\|Ax-Vx\| = \|A-V\|, \end{aligned}$$

где равенство $\|Ax\| = \|Ax-Vx\|$ вытекает из того, что $Vx = 0$ для любого элемента $x \in \mathcal{R}^\perp(V^*) = \mathcal{N}(V)$. Таким образом, при $m < \operatorname{rank} A$ для любого $V \in \mathbb{V}_m$ имеет место неравенство $\|A-V\| \geqslant a_{m+1}$.

Зададим оператор $A_m \in \mathbb{V}_m$ для каждого $x \in \mathcal{R}$ равенством

$$A_m x = \sum_{q=1}^{m} a_q(x, e_q)\tilde{e}_q. \tag{C.33}$$

Тогда при $\|x\| \leqslant 1$ с учётом неубывания сингулярных чисел, формулы (C.28) и неравенства Бесселя

$$\|(A-A_k)x\|^2 = \sum_{q=m+1}^{r} a_q^2(x,e_q)^2 \leqslant a_{m+1}^2\sum_{q=m+1}^{r}(x,e_q)^2 \leqslant a_{m+1}^2\|x\|^2 \leqslant a_{m+1}^2.$$

С другой стороны, $\|(A-A_m)e_{m+1}\| = a_{m+1}$, поэтому для $V = A_m$ имеем

$$\|A-V\| = \sup_{\|x\|\leqslant 1}\|(A-A_m)x\| = a_{m+1}.$$

Формула (C.32) для случая $m < \operatorname{rank} A$ доказана.

Если $\operatorname{rank} A < \infty$ и $m \geqslant \operatorname{rank} A$, то $A \in \mathbb{V}_m$, поэтому минимум в (C.32) достигается при $V = A$ и равен нулю. Теорема доказана.

Если положить, что множество $\mathbb{V}_0$ имеет единственным элементом нулевой оператор, то (C.32) можно считать верной и при $k = 0$ т. е. $a_1 = \|A\| = \|A-V\|$ при $V = 0$.

[8] Будучи оператором конечного ранга, оператор $V \in \mathbb{V}_k$ имеет сопряжённый $V^* \in \mathbb{B}(\widetilde{\mathcal{R}} \mapsto \mathcal{R})$.

Доказанному свойству можно придать следующую геометрическую интерпретацию: величина $\min_{V \in \mathbb{V}_k} \|A - V\|$ есть расстояние от оператора A до операторного множества $\mathbb{V}_k$, а оператор A_k из (C.33) — проекция оператора A на множество $\mathbb{V}_k$.