ТЕОРИЯ ФИЛЬТРА КАЛМАНА

А.И. Матасов

Лаборатория управления и навигации
механико-математический факультет
МГУ им. М.В. Ломоносова

E-mail: alexander.matasov@gmail.com

Часть I.

ГЕОМЕТРИЧЕСКИЕ ОСНОВЫ

ЗАДАЧА ЛИНЕЙНОГО СТОХАСТИЧЕСКОГО ОЦЕНИВАНИЯ

Рассмотрим два случайных вектора $x = x(\omega)$ и $y = y(\omega)$

$$x\colon \Omega \to \mathbb{C}^n, \qquad y\colon \Omega \to \mathbb{C}^p,$$

$$\mathsf{E}\, x = 0, \quad \mathsf{E}\, y = 0, \qquad R_{xy} \triangleq \mathsf{E}\, xy^*, \; R_y \triangleq \mathsf{E}\, yy^* \text{ – известные матрицы.}$$

Линейные оценки вида

$$Ky, \text{ где } K \in \mathbb{C}^{n\times p} \text{ – матричный коэффициент.}$$

Задача минимизации ошибки оценки в *среднеквадратическом смысле:*

$$\mathsf{E}\left[(x - K_0 y)\,(x - K_0 y)^*\right] \le \mathsf{E}\left[(x - Ky)\,(x - Ky)^*\right]$$

(стохастический метод наименьших квадратов).

Свойства

- Покомпонентная оптимальность
- Для любой $W \geq 0, \quad W = W^{\frac{1}{2}} \cdot \left(W^{\frac{1}{2}}\right)^*$

$$\left(W^{\frac{1}{2}}\right)^* \mathsf{E}\left[(x - K_0 y)(x - K_0 y)^*\right] W^{\frac{1}{2}} \leq \left(W^{\frac{1}{2}}\right)^* \mathsf{E}\left[(x - K y)(x - K y)^*\right] W^{\frac{1}{2}} \quad \forall\ K.$$

Следовательно

$$\operatorname{Tr}\left\{\left(W^{\frac{1}{2}}\right)^* \mathsf{E}\left[(x-K_0 y)(x-K_0 y)^*\right] W^{\frac{1}{2}}\right\} \leq \operatorname{Tr}\left\{\left(W^{\frac{1}{2}}\right)^* \mathsf{E}\left[(x-K y)(x-K y)^*\right] W^{\frac{1}{2}}\right\}.$$

Так как $\operatorname{Tr} AA^* = \operatorname{Tr} A^*A$, то

$$\mathsf{E}\left[(x - K_0 y)^* W (x - K_0 y)\right] \leq \mathsf{E}\left[(x - K y)^* W (x - K y)\right] \qquad \forall\ K.$$

Верно и неравенство

$$\det \mathsf{E}\left[(x - K_0 y)(x - K_0 y)^*\right] \leq \det \mathsf{E}\left[(x - K y)(x - K y)^*\right].$$

ОСОБЕННОСТИ КОМПЛЕКСНОЗНАЧНОЙ ФОРМУЛИРОВКИ

Представим векторы x и y в виде их разложения на вещественную и мнимую части:

$$x = x_R + j\, x_I, \qquad y = y_R + j\, y_I, \qquad j^2 = -1.$$

Нетрудно убедиться, что

$$R_y = (R_{y_R} + R_{y_I}) + j\,(R_{y_I,\,y_R} - R_{y_R,\,y_I})$$

и

$$R_{xy} = (R_{x_R,\,y_R} + R_{x_I,\,y_I}) + j\,(R_{x_I,\,y_R} - R_{x_R,\,y_I}).$$

Знание матриц R_{xy} и R_y определяет только некоторые комбинации покомпонентных ковариационных матриц. Поэтому если оперировать с вещественными и мнимыми компонентами x и y и считать известными их ковариационные матрицы, то исследователь будет располагать большей информацией, чем изначально, и оценки получатся более точными.

МАТРИЧНОЗНАЧНЫЕ СКАЛЯРНЫЕ ПРОИЗВЕДЕНИЯ

Линейное пространство. Модуль

Пространство $\mathcal{V}$ с элементами u, v, w, $\ldots$;
ассоциативное кольцо с единицей "скаляров" $\mathcal{S}$ с элементами α, β, $\ldots$;

$$u + v = v + u; \qquad (\alpha + \beta)u = \alpha u + \beta u;$$

$$(u + v) + w = u + (v + w); \qquad (\alpha\beta)u = \alpha(\beta u);$$

$$\alpha(u + v) = \alpha u + \alpha v; \qquad 0 \cdot u = 0, \quad 1 \cdot u = u.$$

Инволюция

Для $\alpha, \beta \in \mathcal{S}$ и $\lambda \in \mathbb{R}^1$ $\quad (\alpha^*)^* = \alpha, \quad (\alpha\beta)^* = \beta^*\alpha^*, \quad (\lambda\alpha)^* = \lambda\alpha^*.$

Скалярное произведение

Билинейная форма на $\mathcal{V}$: $\quad \langle \cdot, \cdot \rangle: \ \mathcal{V} \times \mathcal{V} \to \mathcal{S},$
удовлетворяющая условиям $(u, v, w \in \mathcal{V}, \quad \alpha, \beta \in \mathcal{S})$:

$$\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle;$$

$$\langle v, u \rangle = \langle u, v \rangle^*; \qquad \|u\|^2 = \langle u, u \rangle = 0 \quad \Leftrightarrow \quad u = 0.$$

Основное пространство для линейного оценивания

Тройка $\{\mathcal{V}, \mathcal{S}, \langle\cdot,\cdot\rangle\}$ также будет называться модулем.

Если $\mathcal{S}$ есть $\mathbb{C}$ (или $\mathbb{R}$), то $\{\mathcal{V}, \mathcal{S}, \langle\cdot,\cdot\rangle\}$ – евклидово пространство.

Положим

$$\mathcal{V} = \{Sx + Ky\}\,, \quad \text{где} \;\; S \in \mathbb{C}^{n\times n} \text{ и } K \in \mathbb{C}^{n\times p}.$$

Кольцо $\mathcal{S}$ состоит из квадратных матриц размера $n \times n$; $\;*$ – эрмитово сопряжение.

Иными словами,

$$\mathcal{V} = \left\{ S_1 \begin{pmatrix} x_{(1)} \\ 0_{n-1} \end{pmatrix} + \ldots + S_n \begin{pmatrix} x_{(n)} \\ 0_{n-1} \end{pmatrix} + S_{n+1} \begin{pmatrix} y_{(1)} \\ 0_{n-1} \end{pmatrix} + \ldots + S_{n+p} \begin{pmatrix} y_{(p)} \\ 0_{n-1} \end{pmatrix} \right\}, \quad S_i \in \mathcal{S}.$$

Скалярное произведение двух случайных векторов a и b одинаковой размерности:

$$\langle a, b\rangle = \mathsf{E}\, ab^*; \qquad\qquad \|a\|^2 = \langle a, a\rangle.$$

Базовая задача линейного оценивания принимает вид $\quad \|x - Ky\|^2 \to \min_K$.

Ортогональная проекция

Пусть $\mathcal{L}$ –линейное подпространство $\mathcal{V}$; $\qquad \langle x - \hat{x}, z\rangle = 0 \qquad \forall\, z \in \mathcal{L}$.

РЕШЕНИЕ БАЗОВОЙ ЗАДАЧИ ОЦЕНИВАНИЯ

$\boxed{\text{Ортогональная проекция } x \text{ на } \mathcal{L} = \{Ky\} \text{ решает задачу оптимального оценивания.}}$

Пусть проекция $\hat{x} = K_0 y$ вектора x на $\{Ky\}$ существует:

$$\langle x - \hat{x},\ Ky \rangle = \langle x - K_0 y,\ Ky \rangle = 0 \qquad \forall\ K.$$

Из свойств скалярного произведения:

$$\langle x - Ky,\ x - Ky \rangle = \langle x - K_0 y + (K_0 - K)\, y,\ x - K_0 y + (K_0 - K)\, y \rangle =$$

$$= \|x - K_0\|^2 + \langle x - K_0 y,\ (K_0 - K)\, y \rangle + \langle x - K_0 y,\ (K_0 - K)\, y \rangle^* + \|\,(K_0 - K)\, y\|^2.$$

Второй и третий члены в правой части равны нулю. Тогда

$$\|x - Ky\|^2 = \|x - K_0 y\|^2 + \|\,(K_0 - K)\, y\|^2$$

и

$$\|x - Ky\|^2 \geq \|x - K_0 y\|^2.$$

$\boxed{\text{Никакая другая оценка, кроме ортогональной проекции, не является оптимальной.}}$

Допустим, что оценка $\tilde{K}y$ при $\tilde{K}y \neq K_0 y$ оптимальна. Тогда

$$\|x - \tilde{K}y\|^2 \leq \|x - K_0 y\|^2$$

и

$$0 \geq \|x - \tilde{K}y\|^2 - \|x - K_0 y\|^2 = \|(K_0 - \tilde{K})y\|^2, \qquad \text{что невозможно.}$$

Условие ортогональности эквивалентно более простому условию

$$\langle x - K_0 y,\ y\rangle = 0, \qquad \text{то есть} \qquad \mathsf{E}\, xy^* = K_0\, \mathsf{E}\, yy^*.$$

Система нормальных уравнений: $\qquad R_{xy} = K\, R_y.$

$\boxed{\text{Система нормальных уравнений всегда разрешима.}}$

При $R_y > 0$ $\qquad K_0 = R_{xy} R_y^{-1}, \qquad \hat{x} = R_{xy} R_y^{-1} y.$

Из единственности ортогональной проекции следует, что при наличии двух разных решений K_1 и K_2 оптимальная оценка единственна: $K_1 y = K_2 y$.

$\boxed{\text{Знание о явном соотношении между } x \text{ и } y \text{ не увеличивает точности оценивания.}}$

ПРИМЕРЫ

Пример 1.

Рассмотрим вещественнозначный стационарный случайный процесс $\{y(t)\}$ с нулевым средним и корреляционной функцией

$$\langle y(t), y(t-\tau)\rangle = \mathsf{E}\, y(t)y(t-\tau) = \varrho_y(\tau).$$

Требуется найти наилучшую в среднеквадратическом смысле линейную оценку

$$\int_0^T y(t)\,dt \quad \text{по двум измерениям} \quad y(0) \text{ и } y(T).$$

Р е ш е н и е. Положим $x = \int_0^T y(t)\,dt \in \mathbb{R}$, $y = \operatorname{col}(y(0), y(T)) \in \mathbb{R}^2$.

$$R_{xy} = \int_0^T \varrho_y(t)\,dt\,(1,1), \qquad R_y = \begin{pmatrix} \varrho_y(0) & \varrho_y(T) \\ \varrho_y(T) & \varrho_y(0) \end{pmatrix}.$$

Тогда

$$\hat{x} = \frac{\int_0^T \varrho_y(t)\,dt}{\varrho_y(0) + \varrho_y(T)}\,\bigl(y(0) + y(T)\bigr).$$

Пример 2.

Пусть имеется линейная зависимость вида

$$y = Hx + v,$$

где

- $y \in \mathbb{C}^p$ – измеряемый вектор,
- $H \in \mathbb{C}^{p \times n}$ – заданная матрица,
- $x \in \mathbb{C}^n$ – оцениваемый вектор с нулевым средним и ковариационной матрицей $R_x = \langle x, x \rangle$,
- $v \in \mathbb{C}^p$ – шум измерений с нулевым средним и ковариационной матрицей $R_v = \langle v, v \rangle$,
- x и v некоррелированы друг с другом: $\langle x, v \rangle = 0$.

Требуется найти наилучшую в среднеквадратическом смысле линейную оценку x по y.

Р е ш е н и е. Ясно, что $R_{xy} = R_x H^*$. Положим, что $R_y = HR_xH^* + R_v > 0$. Тогда

$$\hat{x} = R_x H^* \left(HR_xH^* + R_v\right)^{-1} y.$$

Ковариационная матрица ошибки оценки $P_x = \langle x - \hat{x}, x - \hat{x} \rangle$ имеет вид

$$P_x = R_x - R_x H^* \left(HR_xH^* + R_v\right)^{-1} HR_x.$$

СВОЙСТВА ОПТИМАЛЬНЫХ ОЦЕНОК

1. Ортогональная проекция суммы векторов x и z на подпространство $\mathcal{L} = \{Ky\}$ есть сумма ортогональных проекций каждого:

$$\widehat{(x+z)} = \hat{x} + \hat{z}.$$

Действительно, из условия ортогональности

$$\langle (x+z) - (\hat{x} + \hat{z}),\ y\rangle = \langle x - \hat{x},\ y\rangle + \langle z - \hat{z},\ y\rangle = 0.$$

2. Рассмотрим два модуля

$$\mathcal{V} = \{Sx + Ky\},\ \text{где } S \in \mathbb{C}^{n\times n},\ K \in \mathbb{C}^{n\times p} \text{ – произвольные матрицы}$$

и

$$\tilde{\mathcal{V}} = \{\tilde{S}x + \tilde{K}y\},\ \text{где } \tilde{S} \in \mathbb{C}^{m\times n},\ \tilde{K} \in \mathbb{C}^{m\times p} \text{ – произвольные матрицы.}$$

Пусть $L \in \mathbb{C}^{m\times n}$ – заданная матрица.

Ортогональная проекция линейного преобразования x на подпространство $\{\tilde{K}y\}$ есть линейное преобразование ортогональной проекции x на подпространство $\{Ky\}$:

$$\widehat{(Lx)} = L\hat{x}.$$

Действительно, из условия ортогональности $\langle Lx - L\hat{x},\ y\rangle = L\langle x - \hat{x},\ y\rangle = 0.$

3. Ортогональная проекция x на подпространство, порожденное ортогональными векторами y_1 и y_2, есть сумма ортогональных проекций на каждое подпространство:

$$\hat{x}_{\{y\}} = \hat{x}_{\{y_1\}} + \hat{x}_{\{y_2\}}, \qquad y = \mathrm{col}(y_1, y_2), \quad y_1,\, y_2 \in \mathbb{C}^p, \qquad y_1 \perp y_2.$$

Действительно, из линейности $\langle \cdot, \cdot \rangle$ по первому аргументу

$$\langle x - (\hat{x}_{\{y_1\}} + \hat{x}_{\{y_2\}}),\ y_1 \rangle = \langle x - \hat{x}_{\{y_1\}},\ y_1 \rangle - \langle \hat{x}_{\{y_2\}},\ y_1 \rangle = 0.$$

- $\langle x - \hat{x}_{\{y_1\}},\ y_1 \rangle = 0,$ так как $\hat{x}_{\{y_1\}}$ есть ортогональная проекция на подпространство $\{K_1 y_1\}$;
- $\langle \hat{x}_{\{y_2\}},\ y_1 \rangle = 0$ потому, что $\hat{x}_{\{y_2\}} = K_2^0\, y_2$ и следовательно $\langle K_2^0 y_2, y_1 \rangle = K_2^0 \langle y_2, y_1 \rangle = 0$.

Совершенно аналогично,

$$\langle x - (\hat{x}_{\{y_1\}} + \hat{x}_{\{y_2\}}),\ y_2 \rangle = 0.$$

Тогда

$$\langle x - (\hat{x}_{\{y_1\}} + \hat{x}_{\{y_2\}}),\ y \rangle = 0, \qquad y = \mathrm{col}(y_1, y_2).$$

Указанное свойство справедливо для любого конечного числа подвекторов вектора y.

СУЩЕСТВОВАНИЕ И ЕДИНСТВЕННОСТЬ ПРОЕКЦИИ

Рассмотрим модуль

$$\mathcal{V} = \{Sx + Ky\}\,, \qquad \langle x, z\rangle = \mathsf{E}\, xz^*, \qquad x, z \in \mathcal{V},$$

где $S \in \mathbb{C}^{n\times n}$ и $K \in \mathbb{C}^{n\times p}$ – произвольные матрицы.

Пусть $\mathcal{L} = \{Ky\}$ – линейное подпространство в $\mathcal{V}$, порожденное вектором y.

Теорема 1. *Для любого вектора $x \in \mathcal{V}$ существует его ортогональная проекция на $\mathcal{L}$. Более того, она единственна.*

О доказательстве существования проекции

В книге LINEAR ESTIMATION в лемме 4.A.2 на стр. 151 приведено доказательство существования и единственности ортогональной проекции. В нем предполагается, что в модуле существует ортонормированный базис. Приведем контрпример.

Рассмотрим линейное пространство вида

$$\mathcal{V} = \left\{ S \begin{pmatrix} 1 \\ 1 \end{pmatrix} \xi \right\}, \qquad \mathsf{E}\,\xi = 0, \quad \mathsf{E}\,\xi^2 = 1,$$

где S есть произвольная вещественнозначная матрица размером $2{\times}2$, со скалярным произведением $\langle x, y \rangle = \mathsf{E}\,xy^*$. Все элементы этого модуля имеют вырожденные нормы:

$$\|x\|^2 = \langle x, x \rangle = S \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} S^* \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

КОРРЕКТНОЕ ДОКАЗАТЕЛЬСТВО

Пусть $x = \mathrm{col}(x_{(1)}, \ldots, x_{(n)})$, а $y = \mathrm{col}(y_{(1)}, \ldots, y_{(p)})$.

Рассмотрим линейное пространство числовых случайных величин:

$$V = \left\{\alpha_1 x_{(1)} + \ldots, +\alpha_n x_{(n)} + \beta_1 y_{(1)} + \ldots + \beta_p y_{(p)}\right\}, \qquad \alpha_i, \beta_j \in \mathbb{C}^1$$

со скалярным произведением $\langle u, v\rangle = \mathsf{E} v^* u \in \mathbb{C}^1$. Это конечномерное евклидово пространство размерности $N \le n + p$. Поэтому в V существует ортонормированный базис $\{\varrho_1, \ldots, \varrho_N\}$ такой, что $\{\varrho_1, \ldots, \varrho_M\}$ есть базис в $L = \left\{\beta_1 y_{(1)} + \ldots, +\beta_p y_{(p)}\right\}$.

Так как

$$\mathrm{span}\left\{y_{(1)}, \ldots, y_{(p)}\right\} = \mathrm{span}\{\varrho_1, \ldots, \varrho_M\},$$

то подпространство $\mathcal{L}$ пространства $\mathcal{V}$ покомпонентно выражается через базис $\{\varrho_j\}_1^M$:

$$\mathcal{L} = \{Ky\} = \left\{\overset{(1)}{w}\varrho_1 + \ldots + \overset{(M)}{w}\varrho_M\right\}, \qquad \overset{(j)}{w} \in \mathbb{C}^n, \qquad j = 1, \ldots, M.$$

Кроме того,

$$x = \overset{(1)}{\kappa}\,\varrho_1 + \ldots + \overset{(M)}{\kappa}\,\varrho_M + \overset{(M+1)}{\kappa}\,\varrho_{M+1} + \ldots + \overset{(N)}{\kappa}\,\varrho_N, \qquad \overset{(i)}{\kappa} \in \mathbb{C}^n, \quad i = 1, \ldots, N.$$

Положим

$$\hat{x} = \overset{(1)}{\kappa}\,\varrho_1 + \ldots + \overset{(M)}{\kappa}\,\varrho_M \in \mathcal{L}.$$

Ясно, что $x - \hat{x}$ ортогонально к $\mathcal{L}$.

Предположим, что существуют две проекции $\hat{x} \in \mathcal{L}$ и $\hat{z} \in \mathcal{L}$. Тогда

$$\|\hat{x} - \hat{z}\|^2 = \langle \hat{x} - \hat{z}, \hat{x} - \hat{z} \rangle = \langle (x - \hat{z}) - (x - \hat{x}),\ \hat{x} - \hat{z} \rangle = 0 \quad \Leftrightarrow \quad \hat{x} = \hat{z}.$$

АЛГЕБРАИЧЕСКОЕ ДОКАЗАТЕЛЬСТВО

Рассмотрим матрицу $H \in \mathbb{C}^{p\times q}$ и введем образ и ядро оператора H:

$$\mathcal{R}(H) = \{Ha \,|\, a \in \mathbb{C}^q\}, \qquad \mathcal{N}(H) = \{a \in \mathbb{C}^q \,|\, Ha = 0\}.$$

Ортогональным дополнением $L^\perp$ к подпространству $L \subseteq \mathbb{C}^q$ называется подпространство

$$L^\perp = \{a \in \mathbb{C}^q \,|\, a^* c = 0 \text{ для всех } c \in L\}.$$

Лемма 1. *Имеют место следующие свойства:*

$$\mathcal{N}(H) = \mathcal{R}(H^*)^\perp, \quad \mathcal{N}(H^*) = \mathcal{R}(H)^\perp, \quad \mathcal{R}(H) = \mathcal{N}(H^*)^\perp \ \text{ и } \ \mathcal{R}(H^*) = \mathcal{N}(H)^\perp.$$

Доказательство. Докажем первое свойство. С одной стороны,

$$a \in \mathcal{N}(H) \;\Rightarrow\; Ha = 0 \;\Rightarrow\; a^*H^*b = 0 \text{ для всех } b \in \mathbb{C}^p \;\Rightarrow\; a \perp \mathcal{R}(H^*) \;\Rightarrow$$
$$\Rightarrow\; a \in \mathcal{R}(H^*)^\perp \;\Rightarrow\; \mathcal{N}(H) \subseteq \mathcal{R}(H^*)^\perp.$$

С другой стороны,

$$a \in \mathcal{R}(H^*)^\perp \;\Rightarrow\; a \perp \mathcal{R}(H^*) \Rightarrow a^*H^*b = 0 \text{ для всех } b \in \mathbb{C}^p \Rightarrow b^*(Ha) = 0 \text{ для всех } b \in \mathbb{C}^p \Rightarrow$$
$$\Rightarrow\; Ha = 0 \;\Rightarrow\; a \in \mathcal{N}(H) \;\Rightarrow\; \mathcal{R}(H^*)^\perp \subseteq \mathcal{N}(H).$$

Следствие 1. *Справедливы разложения*

$$\mathcal{N}(H) \oplus \mathcal{R}(H^*) = \mathbb{C}^q \quad \text{и} \quad \mathcal{N}(H^*) \oplus \mathcal{R}(H) = \mathbb{C}^p.$$

ТЕОРЕМА 2. *Система нормальных уравнений всегда имеет решение. Причем, если K_1 и K_2 два разных решения, то $K_1 y = K_2 y$.*

Доказательство. Существование решения нормальных уравнений эквивалентно включению $\mathcal{R}(R_{yx}) \subseteq \mathcal{R}(R_y)$. Поскольку $R_{yx} = R^*_{xy}$ и $R_y = R^*_y$, то система принимает вид

$$R_{yx}(\epsilon_1, \ldots, \epsilon_n) = R_y\,(K^*_1, \ldots, K^*_n),$$

где ϵ_i – единичные орты-столбцы в $\mathbb{R}^n$, а K^*_i – столбцы матрицы K^*.

Если эта система имеет решение K^*, то для любого вектора

$$a = a_1\epsilon_1 + \ldots + a_n\epsilon_n, \qquad a_i \in \mathbb{C}^1, \quad i = 1, \ldots, n$$

справедливо равенство

$$R_{yx}a = R_{yx}(a_1\epsilon_1 + \ldots + a_n\epsilon_n) = R_y\,(a_1K^*_1 + \ldots + a_nK^*_n),$$

то есть $\mathcal{R}(R_{yx}) \subseteq \mathcal{R}(R_y)$.

Обратно, если $\mathcal{R}(R_{yx}) \subseteq \mathcal{R}(R_y)$, то существуют K_i^* такие, что

$$R_{yx}\epsilon_i = R_y K_i^*, \qquad i = 1, \ldots, n,$$

то есть нормальные уравнения разрешимы.

Согласно следствию 1,

$$\mathcal{R}(R_{yx}) \oplus \mathcal{N}(R_{xy}) = \mathbb{C}^p = \mathcal{R}(R_y) \oplus \mathcal{N}(R_y).$$

Следовательно

$$\mathcal{R}(R_{yx}) \subseteq \mathcal{R}(R_y) \quad \Leftrightarrow \quad \mathcal{N}(R_y) \subseteq \mathcal{N}(R_{xy}).$$

Пусть $R_y b = 0$ для некоторого $b \in \mathbb{C}^p$. Тогда

$$\mathsf{E}\,(b^* y)(y^* b) = \mathsf{E}\,|y^* b|^2 = 0 \quad \Leftrightarrow \quad y^* b = 0.$$

Поэтому $R_{xy} b = \mathsf{E}\, x(y^* b) = 0$, то есть $b \in \mathcal{N}(R_{xy})$.

Пусть теперь $R_{xy} = K_1 R_y$ и $R_{xy} = K_2 R_y$. Тогда

$$(K_1 - K_2) R_y = 0 \quad \text{и} \quad \mathsf{E}\,(K_1 - K_2) y y^* (K_1 - K_2)^* = \mathsf{E}\,\|(K_1 - K_2) y\|^2 = 0.$$

Поэтому $K_1 y = K_2 y$.

ВЫРОЖДЕННАЯ МАТРИЦА R_y

Псевдообратная матрица

Пусть дана прямоугольная матрица $A \in \mathbb{C}^{p \times q}$.

Определение 1. Псевдообратной матрицей к матрице A называется матрица $A^+ \in \mathbb{C}^{q \times p}$ такая, что выполняются следующие четыре условия:

$$AA^+A = A; \qquad A^+AA^+ = A^+;$$

$$(AA^+)^* = AA^+; \qquad (A^+A)^* = A^+A.$$

Определение 2. Псевдообратной матрицей к A называется матрица $A^+ \in \mathbb{C}^{q \times p}$

$$A^+ = \lim_{\delta \to 0} A^* \left(AA^* + \delta^2 I \right)^{-1} = \lim_{\delta \to 0} \left(A^*A + \delta^2 I \right)^{-1} A^*.$$

Пусть матрица $A \in \mathbb{C}^{p \times q}$ ранга $r \geq 1$ имеет сингулярное разложение

$$A = U \Lambda V^*,$$

где

$$U \in \mathbb{C}^{p \times r}, \qquad \Lambda \in \mathbb{C}^{r \times r}, \qquad V \in \mathbb{C}^{q \times r}, \qquad U^* U = V^* V = I,$$

а матрица

$$\Lambda = \operatorname{diag}\left(\lambda_1, \ldots, \lambda_r\right),$$

в которой $\lambda_1 \geq \lambda_2 \geq \ldots, \geq \lambda_r > 0$ – сингулярные числа, то есть положительные квадратные корни из первых r положительных собственных чисел эрмитовой матрицы AA^* (или A^*A), расположенных в убывающем порядке.

Определение 1.3. Псевдообратной матрицей к A называется матрица $A^+ \in \mathbb{C}^{q \times p}$ вида

$$A^+ = V \Lambda^{-1} U^*.$$

Можно показать, что псевдообратная матрица существует и единственна, а все три определения эквивалентны. Если A обратима, то $A^+ = A^{-1}$.

Решение вырожденных систем линейных уравнений

$$A\chi = b, \qquad A \in \mathbb{C}^{p\times q}, \qquad b \in \mathbb{C}^{p}.$$

Лемма 2. *Система линейных уравнений совместна тогда и только тогда, когда* $AA^{+}b = b$. *Если система линейных уравнений совместна, то ее общее решение имеет вид*

$$\chi = A^{+}b + \left(I - A^{+}A\right) z, \quad \text{где } z \text{ – произвольный элемент из } \mathbb{C}^{q}.$$

Доказательство. Пусть $AA^{+}b = b$, тогда $\chi = A^{+}b$ есть решение.
Пусть теперь система имеет решение: $A\chi = b$. Тогда $AA^{+}b = AA^{+}A\chi = A\chi = b$.

Для любого χ, определяемого формулой для общего решения,

$$A\chi = AA^{+}b + A\left(I - A^{+}A\right) z = b + \left(A - AA^{+}A\right) z = b + 0 = b.$$

Пусть χ – любое решение. Тогда $\chi = A^{+}A\chi + (I - A^{+}A)\,\chi = A^{+}b + (I - A^{+}A)\,\chi$.

Слагаемые в правой части ортогональны друг другу. Действительно,

$$z^{*}\left(I - A^{+}A\right)^{*} A^{+}b = z^{*}\left(I - A^{+}A\right) A^{+}b = z^{*}\left(A^{+} - A^{+}AA^{+}\right) b = 0.$$

Тогда $\|\chi\|^2 = \|A^{+}b\|^2 + \|\,(I - A^{+}A)\,z\|^2$ и $A^{+}b$ есть решение с минимальной нормой.

Решение вырожденной системы нормальных уравнений

Систему нормальных уравнений можно представить в форме

$$R_y\, K^* = R_{yx}, \qquad \text{где } R_y \in \mathbb{C}^{p\times p}, \quad K^* \in \mathbb{C}^{p\times n}, \quad R_{yx} \in \mathbb{C}^{p\times n}.$$

Это ничто иное, как n систем уравнений с матрицей R_y.

Система нормальных уравнений всегда имеет решение. Общий вид решения:

$$K^* = R_y^+ R_{yx} + \left(I - R_y^+ R_y\right) Z, \quad \text{где } Z \text{ – произвольная матрица из } \mathbb{C}^{p\times n}.$$

Так как $(R_y^+)^* = R_y^+$, то

$$K = R_{xy} R_y^+ + Z^* \left(I - R_y R_y^+\right), \quad \text{где } Z^* \text{ – произвольная матрица из } \mathbb{C}^{n\times p}.$$

Поскольку ортогональная проекция единственна (и для любых двух решений $K_1 y = K_2 y$), то ортогональная проекция имеет вид (для простоты можно взять $Z^* = 0$)

$$\hat{x} = R_{xy} R_y^+ y.$$

При невырожденной матрице R_y полученное решение переходит в обычное.

ИННОВАЦИОННАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ

Рассмотрим случайные векторы $y_0, y_1, \ldots, y_N$, $y_i \in \mathbb{C}^p$.

Разумно найти такую последовательность ортогональных векторов $e_0, \ldots, e_N$, что

$$\mathrm{span}\{e_0, \ldots, e_N\} = \mathrm{span}\{y_0, \ldots, y_N\}.$$

Естественный путь построения *инновационной* последовательности:

$$e_{i+1} \stackrel{\Delta}{=} y_{i+1} - \mathrm{Proj.}\{y_{i+1} \,|\, \mathcal{L}_i\} = y_{i+1} - \hat{y}_{i+1}, \qquad i = 0, \ldots, N-1, \qquad e_0 \stackrel{\Delta}{=} y_0.$$

Ясно, что

$$e_i \in \mathcal{L}\{y_0, \ldots, y_i\}, \qquad y_i \in \mathcal{L}\{e_0, \ldots, e_i\};$$

это "физически осуществимые" преобразования.

Стало быть,

$$\mathcal{L}\{y_0, \ldots, y_i\} = \mathcal{L}\{e_0, \ldots, e_i\}, \qquad i = 0, \ldots, N.$$

Из третьего свойства оптимальных оценок получаем, что

$$e_i = y_i - \sum_{j=0}^{i-1} \langle y_i, e_j \rangle \|e_j\|^{-2} e_j, \qquad y_i = e_i + \sum_{j=0}^{i-1} \langle y_i, e_j \rangle \|e_j\|^{-2} e_j.$$

Аналогия с процедурой ортогонализации Грама-Шмидта.

Алгебраический подход

Найти невырожденную матрицу A такую, что

$$e = Ay, \qquad e = \mathrm{col}(e_0, \ldots, e_N),$$

где

$$R_e = \langle e, e \rangle = \langle Ay, Ay \rangle = AR_yA^* \text{ – диагональная матрица.}$$

Тогда

$$R_y = A^{-1}R_e\,A^{-*}, \qquad R_e \text{ – диагональная матрица.}$$

Выбор A неоднозначен: $(N+1)^2$ параметров и $N(N+1)/2$ ограничений. Один выбор доставляется

спектральным разложением эрмитовой матрицы R_y; тогда A унитарна. Однако, пока A не является нижнетреугольной, каждый вектор e_i зависит от всех величин $\{y_0, \ldots, y_N\}$. Поэтому A и A^{-1} должны быть нижнетреугольными:

$$R_y = LDL^*,$$

где $D = R_e$ – диагональная матрица, а L – нижнетреугольная матрица.

Теорема 3 (LDU-разложение). *Пусть квадратная матрица R является сильно регулярной, то есть все ее ведущие миноры ненулевые. Тогда $R = LDU$, где L – нижнетреугольная, U – верхнетреугольная матрица с единичными элементами на диагонали, а D – невырожденная диагональная матрица. Более того, это разложение единственно.*

Следствие 2. *Если R является эрмитовой положительно определенной матрицей, то $R = LDL^*$, где L – нижнетреугольная матрица с единичными элементами на диагонали, а D – диагональная матрица с положительными элементами. Более того, такое разложение единственно.*

Следствие 3. *Если R является эрмитовой положительно определенной матрицей, то $R = LL^*$, где L – нижнетреугольная матрица с положительными элементами на диагонали. Более того, такое разложение единственно.*

Разложение $y_i = e_i + \sum_{j=0}^{i-1} \langle y_i, e_j \rangle \|e_j\|^{-2} e_j$ можно записать в виде

$$y = Le, \qquad \text{где} \qquad L = \begin{pmatrix} I & & & & \\ \langle y_1, e_0 \rangle \|e_0\|^{-2} & I & & & \\ \langle y_2, e_0 \rangle \|e_0\|^{-2} & \langle y_2, e_1 \rangle \|e_1\|^{-2} & I & & \\ \vdots & \vdots & \vdots & \ddots & \\ \langle y_N, e_0 \rangle \|e_0\|^{-2} & \langle y_N, e_1 \rangle \|e_1\|^{-2} & \langle y_N, e_2 \rangle \|e_2\|^{-2} & \dots & I \end{pmatrix}.$$

В силу единственности LDU-разложения, L определяет искомое представление с $D = R_e$.

СЛУЧАЙ НЕНУЛЕВЫХ СРЕДНИХ

Обобщим базовую задачу оценивания:

$$\mathsf{E}x = m_x, \quad \mathsf{E}y = m_y, \qquad \mathsf{E}\left[(x - m_x)(y - m_y)^*\right] = R_{xy}, \quad \mathsf{E}\left[(y - m_y)(y - m_y)^*\right] = R_y.$$

Введем центрированные случайные векторы

$$x^\circ = x - m_x \qquad \text{и} \qquad y^\circ = y - m_y.$$

Интуитивно очевидный способ действий состоит в том, чтобы сначала построить оценку центрированной величины по центрированной, а потом перейти к исходным векторам.

Пусть $\hat{x}^\circ = K_0 y^\circ$ – оптимальная оценка x° по y°. Тогда

$$\hat{x} = m_x + K_0(y - m_y) = K_0 y + (m_x - K_0 m_y).$$

В случае невырожденной матрицы R_y соответствующие оценки примут вид

$$\hat{x}^\circ = R_{xy} R_y^{-1} y^\circ$$

или

$$\hat{x} = m_x + R_{xy} R_y^{-1}(y - m_y) = R_{xy} R_y^{-1} y + \left(m_x - R_{xy} R_y^{-1} m_y\right).$$

Формальные рассуждения

Рассмотрим линейные (аффинные) оцениватели вида

$$\hat{x} = Ky + l, \quad \text{где} \ \ K \in \mathbb{C}^{n\times p}, \quad \text{а} \ \ l \in \mathbb{C}^{n}.$$

Ставится задача нахождения коэффициентов K_0 и l_0, минимизирующих ошибку оценки

$$\mathsf{E}\left[(x - K_0 y - l_0)(x - K_0 y - l_0)^*\right] \le \mathsf{E}\left[(x - Ky - l)(x - Ky - l)^*\right].$$

Нетрудно убедиться, что

$$\mathsf{E}\left[(x - Ky - l)(x - Ky - l)^*\right] = \mathsf{E}\left[(x^\circ - Ky^\circ)(x^\circ - Ky^\circ)^*\right] + \\ + (m_x - Km_y - l)(m_x - Km_y - l)^*.$$

Пусть матричный коэффициент K_0 решает задачу для центрированных величин. Тогда

$$\|x^\circ - Ky^\circ\|^2 = \|x^\circ - K_0 y^\circ\|^2 + \|(K_0 - K)\, y^\circ\|^2$$

и норму ошибку оценки можно представить в форме

$$\|x - Ky - l\|^2 = \|x^\circ - K_0 y^\circ\|^2 + \|(K_0 - K)\, y^\circ\|^2 + \|m_x - Km_y - l\|^2.$$

Положим $l_0 = m_x - K_0 m_y$. При таком выборе l_0 имеем

$$\|x - K_0 y - l_0\|^2 = \|x^\circ - K_0 y^\circ\|^2$$

и тогда норма ошибки оценки принимает вид

$$\|x - Ky - l\|^2 = \|x - K_0 y - l_0\|^2 + \|\left(K_0 - K\right) y^\circ\|^2 + \|m_x - K m_y - l\|^2.$$

Отсюда вытекает, что пара $\{K_0, l_0\}$ решает поставленную задачу.
При этом оценка $K_0 y + l_0$ совпадает с оценкой, полученной из интуитивных соображений.

Допустим, что какая-то другая оценка $\tilde{K} y + \tilde{l}$ также является оптимальной. Тогда

$$0 \geq \|x - \tilde{K} y - \tilde{l}\|^2 - \|x - K_0 y - l_0\|^2 = \|(K_0 - \tilde{K}) y^\circ\|^2 + \|m_x - \tilde{K} m_y - \tilde{l}\|^2.$$

Последнее неравенство возможно, только если $K_0 y^\circ = \tilde{K} y^\circ$ и $\tilde{l} = m_x - \tilde{K} m_y$.
Но в этом случае

$$K_0 y + l_0 = K_0 y^\circ + m_x \quad \text{и} \quad \tilde{K} y + \tilde{l} = \tilde{K} y^\circ + m_x,$$

а это означает, что оценки совпадают.