МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**Федеральное государственное автономное образовательное учреждение высшего образования**

**«Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского»**

**Институт информационных технологий, математики и механики**

**Направление подготовки: «Прикладная математика и информатика»**

**Р.С. Бирюков, В.А. Паранин**

# Методические указания по курсу теория управления

Учебно-методическое пособие

*Рекомендовано автором для успешной сдачи экзамена*

Нижний Новгород

2022

## Предисловие

В данном методическом пособии будут разобраны основные вопросы восьмого семестра курса теории управления. Материалы подготовлены на основе лекций Бирюкова Руслана Сергеевича и практик Кадиной Елены Юрьевны, преподавателей данного курса в Нижегородском государственном университете имени Лобачевского.

Не стоит переживать по поводу точности предоставленных ими материалов, почти все они были проверены и исправлены, насколько это было возможно. А если вы нашли ошибку, сообщите автору, чтобы он поправил.

Основными темами данной работы будут управление стохастическими дискретными системами и линейно-матричные неравенства. Мы рассмотрим задачу линейно-квадратичного стохастического управления при полной информации о векторе состояний, немного коснёмся теории оценивания и выведем фильтр Калмана для задачи линейно-квадратичного стохастического управления при неполной информации о векторе состояний. Определим новые понятия устойчивости, решим задачу модального управления относительно $LMI$-области и задачу робастной устойчивости. В общем, интересного будет много. Держитесь крепче, мы взлетаем.

# Оглавление

Предисловие .......... 2
Немного о многомерных случайных величинах .......... 4
Стохастическая оптимальная линейная система при полной информации о состоянии .......... 6
Постановка задачи оценивания. Метод наименьших квадратов и его рекуррентная версия .......... 11
Построение оптимальной оценки по методу минимума дисперсии для одношагового процесса .......... 14
Условное математическое ожидание и ковариационная матрица для гауссовских векторов .......... 18
Рекуррентное оценивание по методу минимума дисперсии .......... 23
Постановка задачи оптимальной фильтрации. Фильтр Калмана .......... 25
Стохастическая оптимальная линейная система при неполной информации о состоянии. Задача линейно-квадратичного гауссовского управления. Принцип разделения. Общая структура решения .......... 27
Линейно-матричные неравенства. Связь с уравнением Ляпунова .......... 29
$LMI$-область: определение и свойства. Понятие $D$-устойчивости линейной системы. Теорема об устойчивости относительно $LMI$-области .......... 34
Задача модального управления относительно $LMI$-области .......... 40
Робастная устойчивость. Теорема Харитонова. Теорема Цыпкина-Поляка. Синтез задачи робастной стабилизации при помощи $LMI$ .......... 42
Приложение .......... 46
Литература и другие ресурсы .......... 47

## Немного о многомерных случайных величинах

На протяжении всего курса нам часто придётся сталкиваться с многомерными случайными величинами, поэтому необходимо вспомнить, ну или узнать, их основные характеристики и свойства.

*Определение.* Многомерной случайной величиной называют измеримое отображение $\xi: \Omega \to \mathbb{R}^n$, где $\Omega$ – множество исходов, $\mathbb{R}^n$ – вещественный $n$-мерный вектор.

У многомерных случайных величин есть несколько важных характеристик, которые будут нас интересовать – математическое ожидание, дисперсия, ковариация и корреляция.

Пусть $x$ и $y$ – $n$-мерные случайные вектора:

$$x = rand(x_1 \quad x_2 \quad \cdots \quad x_n)^T, \qquad y = rand(y_1 \quad y_2 \quad \cdots \quad y_n)^T$$

1. Математическое ожидание

$$M\{x\} = (\bar{x}_1 \quad \bar{x}_2 \quad \cdots \quad \bar{x}_n)^T = \bar{x}$$

2. Ковариация случайных векторов (ковариационная матрица)

$$K_{xy} = M\{(x-\bar{x})(y-\bar{y})^T\} = \begin{pmatrix} cov(x_1, y_1) & \dots & cov(x_1, y_m) \\ \dots & \dots & \dots \\ cov(x_n, y_1) & \dots & cov(x_n, y_m) \end{pmatrix} \in \mathbb{R}^{n\times m}$$

При этом, если это один и тот же вектор, то $K_{xx} = K_{xx}^T \geq 0$

3. Дисперсия

$$Dx = M\{(x-\bar{x})^T(x-\bar{x})\} = tr(K_{xx})$$

4. Обобщённая дисперсия

$$D_{\text{об}}x = \det K_{xx}$$

5. Корреляция

$$R_{xy} = M\{xy^T\}$$

*Определение.* Два случайных вектора $x$ и $y$ не коррелированы, если

$$M\{xy^T\} = M\{x\}M\{y^T\}$$

Отметим важное *свойство*:

$$K_{xy} = R_{xy} - M\{x\}M\{y^T\}$$

при этом, если случайные величины не коррелированы, то матрица ковариации будет нулевой, то есть $K_{xy} = 0$.

В целом, это всё что нам нужно знать о характеристиках и свойствах случайных величин. Однако мы ещё не коснулись самого главного, не коснулись того, что будет преследовать нас на протяжении всего курса. Куда бы мы не бежали от этого, оно нас настигнет. Итак, распределение гаусса (или нормальное распределение).

$$x \sim \mathcal{N}(a, \Sigma) \Leftrightarrow f(x) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det\Sigma}} e^{-\frac{1}{2}(x-a)^T\Sigma^{-1}(x-a)}, \qquad x \in \mathbb{R}^n$$

при этом $\Sigma = \Sigma^T > 0$ (матрица $\Sigma \in \mathbb{R}^{n\times n}$ симметричная положительно определенная).

Вообще говоря, эта матрица не всегда будет положительно определенной, она может быть неотрицательно определена, но в таком случае обратной существовать не будет и нельзя будет так удобно представить функцию распределения. Нас такой вариант не интересует, поэтому мы будем рассматривать только положительно определённые матрицы.

Отметим важные свойства данного распределения

a) $a$ – мат. ожидание

$$M\{x\} = a$$

b) $\Sigma$ – ковариационная матрица

$$K_{xx} = M\{(x-\bar{x})(x-\bar{x})^T\} = \Sigma$$

Этих сведений должно быть достаточно, для понимания происходящего в дальнейшем, но если вы чувствуете, что у вас остались пробелы или вы хотите чуть подробнее разобраться в теме, то можете изучить литературу по этой теме, которая размещена в источниках.

## Стохастическая оптимальная линейная система при полной информации о состоянии

Итак, мы будем рассматривать линейное дискретное динамическое уравнение с полной информацией о состоянии:

$$x_{k+1} = Ax_k + Bu_k + v_k, \qquad k = 0, \dots, N-1$$

где $v_k$ – случайная помеха

$$v_k \sim \mathcal{N}(0, V_k), \qquad x_0 \sim \mathcal{N}(\bar{x}_0, X_0)$$

Случайные помехи на разных шагах не коррелированы, следовательно

$$M\left\{(v_k - 0)\left(v_j - 0\right)^T\right\} = \begin{cases} 0, k \neq j \\ V_k, k = j \end{cases}$$

также случайные помехи не коррелированы с начальным значением состояния системы, а значит

$$M\{(v_k - 0)(x_0 - \bar{x}_0)^T\} = 0$$

Наша цель в заданной задаче – минимизировать функционал:

$$J = M\left\{x_N^T S x_N + \sum_{i=0}^{N-1}\left(x_i^T Q x_i + u_i^T R u_i\right)\right\} \to \min_{u_0, \dots, u_{N-1}}$$

$$S = S^T \geq 0, \qquad Q = Q^T \geq 0\,, \qquad R = R^T > 0$$

Для решения поставленной задачи воспользуемся принципом динамического программирования и запишем функции Беллмана от конца.

Однако перед этим необходимо понять один момент. Все состояния системы $(x_k)$ получаются одно за другим по цепочке

$$x_0 \to x_1 \to x_2 \to \cdots \to x_k \to \cdots$$

При этом, всё это случайные величины, а значит каждое следующее состояние зависит от всех предыдущих. Эту зависимость можно представить, как функцию текущего состояния от предыдущих:

$$x_k = x_k(x_0, x_1, \dots, x_{k-1})$$

И это нужно учесть при составлении функции Беллмана. В итоге у нас будет не просто мат. ожидание, а условное мат. ожидание, так как нам нужно учесть, все предыдущие значения.

$$V_k(x_k) = \min_{u_k, \dots, u_{N-1}} M\left\{x_N^T S x_N + \sum_{i=k}^{N-1}\left(x_i^T Q x_i + u_i^T R u_i\right) | x_k\right\}$$

*Примечание.* Подробнее об условном мат. ожидании можно прочитать в источниках.

Ну-с приступит к преобразованиям.

1) Вынесем из-под знака суммы первое слагаемое

$$V_k(x_k) = \min_{u_k,\dots,u_{N-1}} M\left\{x_N^T S x_N + \sum_{i=k+1}^{N-1}\left(x_i^T Q x_i + u_i^T R u_i\right) + x_k^T Q x_k + u_k^T R u_k | x_k\right\}$$

2) Применим лемму о разделимости, не забыв про условное мат ожидание

$$V_k(x_k) = \min_{u_k} M\left\{\min_{u_{k+1},\dots,u_{N-1}} M\left\{x_N^T S x_N + \sum_{i=k+1}^{N-1}\left(x_i^T Q x_i + u_i^T R u_i\right) | x_{k+1}\right\} + x_k^T Q x_k + u_k^T R u_k | x_k\right\}$$

3) Обратим внимание, что внутренний минимум также является функцией Беллмана

$$V_k(x_k) = \min_{u_k} M\{V_{k+1}(x_{k+1}) \ + x_k^T Q x_k + u_k^T R u_k | x_k\}$$

Теперь сделаем допущение о том, что функцию Беллмана можно представить в виде квадрики, а именно:

$$V_k(x_k) = x_k^T X_k x_k + 2\lambda_k^T x_k + \chi_k$$

Почему так можно сделать? Давайте посмотрим на функцию Беллмана на последнем шаге:

$$V_N(x_N) = x_N^T S x_N$$

Отсюда мы и можем сделать предположение о том, что функцию Беллмана можно представить в виде квадрики, причём мы берём именно полную квадрику, потому что не можем сразу сказать, нужны будут нам остальные слагаемые или нет, а если не можем, значит надо их учесть.

Осталось всё подставить и посчитать. Погнали.

$$x_k^T X_k x_k + 2\lambda_k^T x_k + \chi_k = \min_{u_k} M\{x_{k+1}^T X_{k+1} x_{k+1} + 2\lambda_{k+1}^T x_{k+1} + \chi_{k+1} \ + x_k^T Q x_k + u_k^T R u_k | x_k\} = \min_{u_k} f(u_k)$$

1) Преобразуем выражение в правой части, подставив вместо $x_{k+1}$ уравнение динамики

$$f(u_k) = M\{(Ax_k + Bu_k + v_k)^T X_{k+1}(Ax_k + Bu_k + v_k) + 2\lambda_{k+1}^T(Ax_k + Bu_k + v_k) + \chi_{k+1} + x_k^T Q x_k + u_k^T R u_k | x_k\}$$

2) Раскроем скобки

$$f(u_k) = M\{(Ax_k + Bu_k)^T X_{k+1}(Ax_k + Bu_k) + 2(Ax_k + Bu_k)^T X_{k+1} v_k + v_k^T X_{k+1} v_k + 2\lambda_{k+1}^T(Ax_k + Bu_k + v_k) + \chi_{k+1} + x_k^T Q x_k + u_k^T R u_k | x_k\}$$

Почему именно так? Потому что пока так удобнее.

3) Раскроем мат. ожидание

- Все слагаемые без члена $v_k$ останутся такими же
- Слагаемые, где $v_k$ встречается один раз, занулятся, так как $M\{v_k\} = 0$

Осталось одно слагаемое: $M\{v_k^T X_{k+1} v_k\}$. Чтобы раскрыть это слагаемое воспользуемся леммой о мат. ожидании квадратичной формы:

$$M\{v_k^T X_{k+1} v_k\} = M\{v_k^T\} X_{k+1} M\{v_k\} + tr\{X_{k+1} V_k\} = tr\{X_{k+1} V_k\}$$

Подставив всё это, получим следующий результат.

$$f(u_k) = (Ax_k + Bu_k)^T X_{k+1}(Ax_k + Bu_k) + 2\lambda_{k+1}^T(Ax_k + Bu_k) + \chi_{k+1} + x_k^T Q x_k + u_k^T R u_k + tr\{X_{k+1} V_k\}$$

Найдём экстремум этой функции и убедимся, что это точка минимума.

1) Ищем градиент

$$\nabla f = 2B^T X_{k+1}(Ax_k + Bu_k) + 2B^T \lambda_{k+1} + 2Ru_k = 0$$

2) Ищем стационарную точку

$$B^T X_{k+1} A x_k + B^T \lambda_{k+1} + (R + B^T X_{k+1} B) u_k = 0$$

$$u_k^* = -(R + B^T X_{k+1} B)^{-1} B^T (X_{k+1} A x_k + \lambda_{k+1})$$

3) Проверяем, что это точка минимума

$$\nabla^2 f = 2B^T X_{k+1} B + 2R^T > 0, \qquad (B^T X_{k+1} B \geq 0, R > 0)$$

Гессиан положительно определён, значит $u_k^*$ – точка минимума.

Мы уже на финишной прямой, осталось лишь подставить полученное управление в уравнение для функции Беллмана:

$$x_k^T X_k x_k + 2\lambda_k^T x_k + \chi_k = f(u_k^*)$$

Но перед этим сделаем небольшую замену:

$$W = R + B^T X_{k+1} B \Rightarrow u_k^* = -W^{-1} B^T (X_{k+1} A x_k + \lambda_{k+1})$$

1) Раскроем все слагаемые с управлением в функции $f(u_k^*)$

Ох, что нас ждёт…

$$
\begin{aligned}
Ax_k + Bu_k^* &= Ax_k - BW^{-1}B^T(X_{k+1}Ax_k + \lambda_{k+1}) \\
&= (A - BW^{-1}B^T X_{k+1}A)x_k - BW^{-1}B^T\lambda_{k+1}
\end{aligned}
$$

a) $(Ax_k + Bu_k^*)^T X_{k+1}(Ax_k + Bu_k^*)$

$$
\begin{aligned}
(Ax_k + Bu_k^*)^T X_{k+1}(Ax_k + Bu_k^*) \\
&= x_k^T(A^T - A^T X_{k+1}BW^{-1}B^T)X_{k+1}(A - BW^{-1}B^T X_{k+1}A)x_k \\
&- 2\lambda_{k+1}^T BW^{-1}B^T X_{k+1}(A - BW^{-1}B^T X_{k+1}A)x_k \\
&+ \lambda_{k+1}^T BW^{-1}B^T X_{k+1}BW^{-1}B^T\lambda_{k+1}
\end{aligned}
$$

b) $2\lambda_{k+1}^T(Ax_k + Bu_k^*)$

$$
2\lambda_{k+1}^T(A - BW^{-1}B^T X_{k+1}A)x_k - 2\lambda_{k+1}^T BW^{-1}B^T\lambda_{k+1}
$$

c) ${u_k^*}^T R u_k^*$

$$
\begin{aligned}
{u_k^*}^T R u_k^* &= (\lambda_{k+1}^T + x_k^T A^T X_{k+1})BW^{-1}RW^{-1}B^T(X_{k+1}Ax_k + \lambda_{k+1}) \\
&= x_k^T A^T X_{k+1}BW^{-1}RW^{-1}B^T X_{k+1}Ax_k \\
&+ 2\lambda_{k+1}^T BW^{-1}RW^{-1}B^T X_{k+1}Ax_k + \lambda_{k+1}^T BW^{-1}RW^{-1}B^T\lambda_{k+1}
\end{aligned}
$$

2) Сгруппируем слагаемые в функции $f(u_k^*)$

$$
\begin{aligned}
f(u_k^*) = \chi_{k+1} &+ tr\{X_{k+1}V_k\} \\
&+ x_k^T\{(A^T - A^T X_{k+1}BW^{-1}B^T)X_{k+1}(A - BW^{-1}B^T X_{k+1}A) \\
&+ A^T X_{k+1}BW^{-1}RW^{-1}B^T X_{k+1}A + Q\}x_k \\
&+ \lambda_{k+1}^T\{BW^{-1}B^T X_{k+1}BW^{-1}B^T + BW^{-1}RW^{-1}B^T \\
&- 2BW^{-1}B^T\}\lambda_{k+1} \\
&+ 2\lambda_{k+1}^T\{BW^{-1}RW^{-1}B^T X_{k+1}A + A - BW^{-1}B^T X_{k+1}A \\
&- BW^{-1}B^T X_{k+1}(A - BW^{-1}B^T X_{k+1}A)\}x_k
\end{aligned}
$$

3) Упростим выражения в скобках

a) $x_k^T(\dots)x_k$

$$
\begin{aligned}
(A^T - A^T X_{k+1}BW^{-1}B^T)X_{k+1}(A - BW^{-1}B^T X_{k+1}A) \\
&+ A^T X_{k+1}BW^{-1}RW^{-1}B^T X_{k+1}A + Q \\
&= Q + A^T X_{k+1}A + A^T X_{k+1}BW^{-1}B^T X_{k+1}BW^{-1}B^T X_{k+1}A \\
&- 2A^T X_{k+1}BW^{-1}B^T X_{k+1}A + A^T X_{k+1}BW^{-1}RW^{-1}B^T X_{k+1}A \\
&= Q + A^T X_{k+1}A + A^T X_{k+1}(\dots)X_{k+1}A
\end{aligned}
$$

$$
\dots = BW^{-1}B^T X_{k+1}BW^{-1}B^T - 2BW^{-1}B^T + BW^{-1}RW^{-1}B^T = -BW^{-1}B^T
$$

После подстановки получим окончательное выражение для этой скобки

$$
Q + A^T X_{k+1}A - A^T X_{k+1}BW^{-1}B^T X_{k+1}A
$$

b) $\lambda_{k+1}^T(\dots)\lambda_{k+1}$

Выражение для этой скобки нам уже известно (см. чуть выше):

$$-BW^{-1}B^T$$

c) $\lambda_{k+1}^T(\dots)x_k$

$$BW^{-1}RW^{-1}B^TX_{k+1}A + A - BW^{-1}B^TX_{k+1}A - BW^{-1}B^TX_{k+1}(A - BW^{-1}B^TX_{k+1}A) = A + (\dots)X_{k+1}A$$

$$\dots = BW^{-1}RW^{-1}B^T - 2BW^{-1}B^T + BW^{-1}B^TX_{k+1}BW^{-1}B^T = -BW^{-1}B^T$$

Итого:

$$A - BW^{-1}B^TX_{k+1}A$$

Ура! Мы справились со всеми слагаемыми, осталось записать всё в одном выражении:

$$\begin{aligned}x_k^TX_kx_k + 2\lambda_k^Tx_k + \chi_k &= x_k^T(Q + A^TX_{k+1}A - A^TX_{k+1}BW^{-1}B^TX_{k+1}A)x_k \\ &- \lambda_{k+1}^TBW^{-1}B^T\lambda_{k+1} + 2\lambda_{k+1}^T(A - BW^{-1}B^TX_{k+1}A)x_k \\ &+ tr\{X_{k+1}V_k\} + \chi_{k+1}\end{aligned}$$

Вот и всё. Приводим подобные слагаемые:

$$\begin{cases}X_k = Q + A^TX_{k+1}A - A^TX_{k+1}BW^{-1}B^TX_{k+1}A \\ \lambda_k^T = \lambda_{k+1}^T(A - BW^{-1}B^TX_{k+1}A) \\ \chi_k = \chi_{k+1} + tr\{X_{k+1}V_k\} - \lambda_{k+1}^TBW^{-1}B^T\lambda_{k+1}\end{cases}$$

Функция Беллмана на $N$-ом шаге имеет такие коэффициенты:

$$X_N = S, \chi_N = 0, \lambda_N = 0$$

Из второго уравнения и начальных условий следует, что $\forall k$: $\lambda_k = 0$, а значит формулы примут более простой вид, ну и не забудем, что $W = R + B^TX_{k+1}B$:

$$\begin{cases}X_k = Q + A^TX_{k+1}A - A^TX_{k+1}B(R + B^TX_{k+1}B)^{-1}B^TX_{k+1}A \\ \chi_k = \chi_{k+1} + tr\{X_{k+1}V_k\}\end{cases}$$

Мы почти закончили, осталось лишь найти минимум исследуемого функционала:

$$J = V_0(x_0) = x_0^TX_0x_0 + \chi_0$$

Ну и не забудем про стабилизирующее управление:

$$u_k^* = -(R + B^TX_{k+1}B)^{-1}B^TX_{k+1}Ax_k$$

Это была только первая задача. Дальше – больше!

## Постановка задачи оценивания. Метод наименьших квадратов и его рекуррентная версия

Прошлую задачу мы рассматривали в предположении, что нам полностью известно состояние системы, но, к сожалению, так бывает не всегда, тем не менее нам всё-таки нужно знать состояние системы, пусть и примерно. Чтобы найти это "примерно", мы будем применять различного рода оценки. И первой оценкой, которую мы рассмотрим, будет оценка по методу наименьших квадратов.

Итак, давайте чётко сформулируем задачу.

Мы проводим серию из $N$ испытаний, в которых фиксируем выход системы $y_k$, при этом нам известна связь между выходом системы и её состоянием:

$$y_k = Cx + \omega_k, \qquad x \in \mathbb{R}^n;\ y, \omega \in \mathbb{R}^m$$

где $\omega_k$ – нормально распределённая случайная помеха не коррелированная на разных шагах:

$$\omega_k \sim \mathcal{N}(0, W_k), \qquad M\left\{(\omega_k - 0)\left(\omega_j - 0\right)^T\right\} = \begin{cases} 0, k \neq j \\ W_k, k = j \end{cases}$$

Попробуем применить метод наименьших квадратов для решения этой задачи. Суть этого метода заключается в минимизации квадрата нормы невязки:

$$\left\|\hat{y} - \hat{C}\hat{x}\right\|_2^2 \to \min_{\hat{x}}$$

$\hat{x}$ – оценка состояния системы, $\hat{y} = (y_1^T\ y_2^T \dots y_N^T)^T$ – вектор выходов, $\hat{C} = (C^T\ C^T \dots C^T)^T$ – матрица матриц наблюдения.

Вы можете задать резонный вопрос, куда делась помеха. У меня есть ответ на это: так как помеха случайная величина, то при её учёте в таком выражении результат также будет случайной величиной, от чего мы вообще-то хотели избавиться, вот так вот. Вы можете сказать, что можно взять мат. ожидание от помехи и решать с ним. Что ж, можно, но оно равно нулю, так что получится тоже самое выражение. Однако, если бы помеха не была центрированной, то мы бы это учли, скорее всего. (^_^)

Что ж, мы немного отвлеклись, пора бы и найти уже оценку. Для этого раскроем норму, найдём точку экстремума для полученного выражения и убедимся, что это точка минимума.

1) Раскрываем норму

$$\left\|\hat{y} - \hat{C}\hat{x}\right\|_2^2 = \left(\hat{y} - \hat{C}\hat{x}\right)^2 = \hat{y}^T\hat{y} - 2\hat{x}^T\hat{C}^T\hat{y} + \hat{x}^T\hat{C}^T\hat{C}\hat{x} = f(\hat{x})$$

2) Ищем стационарную точку

$$\nabla f(\hat{x}) = -2\hat{C}^T\hat{y} + 2\hat{C}^T\hat{C}\hat{x} = 0 \Rightarrow \hat{x}^* = \left(\hat{C}^T\hat{C}\right)^{-1}\hat{C}^T\hat{y}$$

3) Проверяем, что это точка минимума

$$\nabla^2 f(\hat{x}) = \hat{C}^T\hat{C} > 0 \Rightarrow \hat{x}^* - \text{т.}\min$$

Получена первая оценка для состояния системы по выходу:

$$\hat{x}^* = \left(\hat{C}^T\hat{C}\right)^{-1}\hat{C}^T\hat{y}$$

Что ж, здорово, конечно, но она не очень удобная, потому что с каждым новым измерением нам придётся пересчитывать всё заново, а мы ещё и помеху никак не учитываем.

Давайте разбираться с проблемами постепенно. Сейчас мы построим удобную рекуррентную оценку, а учётом помехи займёмся в следующий раз.

Допустим мы знаем оценку на $k$-ом шаге:

$$\hat{x}_k = \hat{x}_k(y_1, y_2, \dots, y_k)$$

это так называемая априорная оценка. И мы хотим узнать оценку на следующем шаге с учётом знаний о предыдущей оценке и новом выходе:

$$\hat{x}_{k+1} = \hat{x}_{k+1}(y_1, y_2, \dots, y_{k+1}) = f(\hat{x}_k, y_{k+1})$$

Для решения данной задачи введём новый функционал:

$$G_k(\hat{x}_k) = \sum_{i=1}^{k}(y_i - C\hat{x}_k)^2$$

Этот функционал использует ту же самую идею метода минимума квадратов только не для всех выходов сразу, а для каждого отдельно.

Сделаем с ним всё то же самое, что и раньше: градиент, стационарная точка, проверка на минимум. Поехали

1) Ищем градиент и выражаем стационарную точку

$$\nabla G_k = -2C^T\sum_{i=1}^{k}(y_i - C\hat{x}_k) = 0 \Leftrightarrow -2C^T\sum_{i=1}^{k}y_i + 2kC^TC\hat{x}_k = 0$$

$$\hat{x}_k^* = \frac{1}{k}(C^TC)^{-1}C^T\sum_{i=1}^{k}y_i$$

2) Проверяем на минимум

$$\nabla^2 G_k = 2kC^TC > 0 \Rightarrow \hat{x}_k^* - \text{т.} \min$$

Осталось выразить связь между $k$-ым шагом и $k+1$-ым.

$$\begin{aligned}\hat{x}_{k+1} &= \frac{1}{k+1}(C^TC)^{-1}C^T\sum_{i=1}^{k+1} y_i \\ &= \frac{k}{k+1}\frac{1}{k}(C^TC)^{-1}C^T\sum_{i=1}^{k} y_i + \frac{1}{k+1}(C^TC)^{-1}C^Ty_{k+1} \\ &= \frac{k}{k+1}\hat{x}_k + \frac{1}{k+1}(C^TC)^{-1}C^Ty_{k+1}\end{aligned}$$

Немножко преобразуем первое слагаемое:

$$\frac{k}{k+1}\hat{x}_k = \left(1 - \frac{1}{k+1}\right)\hat{x}_k = \left(I - \frac{(C^TC)^{-1}C^TC}{k+1}\right)\hat{x}_k$$

и получим красивый итоговый ответ:

$$\hat{x}_{k+1} = \hat{x}_k + K_{k+1}(y_{k+1} - C\hat{x}_k), \qquad K_{k+1} = \frac{1}{k+1}(C^TC)^{-1}C^T$$

*Важно.* Из метода построения оценки следует, что $\hat{x}_0 = 0$

*Примечание.* В обеих формулах есть множитель $(C^TC)^{-1}$, поэтому эти формулы не всегда можно применить и поэтому, когда мы писали, что данная матрица положительно определена, мы исключали случай с её вырожденностью. У матрицы $(C^TC)^{-1}$ будет существовать обратная тогда и только тогда, когда матрица $C$ имеет полный столбцовый ранг.

Напоследок давайте разберём небольшой пример.

Пример.

$$y = 2x + \omega, \qquad y_1 = 2, y_2 = 1.8$$

a) Одношаговый процесс

$$\hat{C} = (2 \quad 2)^T, \qquad \hat{y} = (2 \quad 1.8)^T$$

$$\hat{x} = 8^{-1}(2 \quad 2)\begin{pmatrix}2.0\\1.8\end{pmatrix} = 0.95$$

b) Многошаговый процесс

$$\hat{x}_1 = 0 + 1 \cdot 0.25 \cdot 2 \cdot (2 - 0) = 1, \hat{x}_2 = 1 + 0.5 \cdot 0.25 \cdot 2 \cdot (1.8 - 2 \cdot 1) = 0.95$$

На этом наши приключения с методом наименьших квадратов заканчиваются. Поблагодарим его за то, что он такой простой и лояльный к неискушённому читателю. Дальше метод минимума дисперсии, крепитесь.

## Построение оптимальной оценки по методу минимума дисперсии для одношагового процесса

Мы продолжаем оценивать неизвестное состояние системы по известному выходу. Давайте вспомним постановку задачи.

Проводится серия испытаний, в которых фиксируется выход системы $y_k$, при этом нам известна связь между выходом системы и её состоянием:

$$y_k = Cx + \omega_k, \qquad x \in \mathbb{R}^n;\ y, \omega \in \mathbb{R}^m;\ C \in \mathbb{R}^{m \times n}$$

где $\omega_k$ – нормально распределённая случайная помеха не коррелированная на разных шагах:

$$\omega_k \sim \mathcal{N}(0, W_k), \qquad M\left\{(\omega_k - 0)\left(\omega_j - 0\right)^T\right\} = \begin{cases} 0, k \neq j \\ W_k, k = j \end{cases}$$

Если вы внимательно читали предыдущую главу, то помните, что метод наименьших квадратов не позволял нам оценить влияние случайной помехи на нашу оценку, что мы и исправим в этой главе. Для этой цели мы будем применять новый метод оценивания – метод минимума дисперсии.

Но перед решением немного уточним задачу. До этого нам было неважно, что представляет из себя $x$, то есть случайная это величина или детерминированная, сейчас же мы определимся, что это нормально распределённая случайная величина:

$$x \sim \mathcal{N}(\bar{x}, X)$$

В связи с этим введём также дополнительное условие о том, что помехи не коррелированы с состоянием:

$$M\{(\omega_k - 0)(x - \bar{x})^T\} = 0$$

Мы будем искать одношаговую оценку, следовательно, нам не нужна информация о каждом отдельном шаге, и мы можем записать задачу проще после некоторых обозначений и переобозначений.

$$y = (y_1^T\ y_2^T \dots y_N^T)^T, \qquad C = (C^T\ C^T \dots\ C^T)^T, \qquad \omega = (\omega_1^T\ \omega_2^T \dots \omega_N^T)^T$$

$$M\{\omega\} = 0, \qquad M\{(\omega - 0)(\omega - 0)^T\} = W, \qquad M\{(\omega - 0)(x - \bar{x})^T\} = 0$$

В итоге имеют такую задачу:

$$y = Cx + \omega, \qquad \omega \sim \mathcal{N}(0, W), \qquad x \sim \mathcal{N}(\bar{x}, X)$$

$$M\{(\omega - 0)(x - \bar{x})^T\} = 0$$

Задачу мы получили, пора бы и начать её решать, а то что-то мы засиделись.

Как понятно из названия метода, для получения оценки нам нужно найти минимум дисперсии, точнее сказать минимум среднеквадратичной ошибки, аргумент этого минимума и будет нашей оценкой:

$$\hat{x} = \arg\min_{\hat{x}} M\{(x-\hat{x})^T(x-\hat{x})|y\}$$

Ну и так как мы ищем этот минимум в зависимости от выхода $y$, необходимо учесть это в нашем функционале, что мы и сделали.

Мы будем искать решение с конкретно заданным видом оценки:

$$\hat{x} = Ly, \qquad L \in \mathbb{R}^{n\times m}$$

Решать задачу в исходном виде, то есть без дополнительного предположения о виде оценки, мы тоже будем, но немного позднее. Пока ограничимся этим. Должен заметить, что это очень сомнительный вариант, но на экзамене спросят, так что уж постарайтесь.

Ну что ж, давайте решать. Примерный план действий такой: распишем выражение дисперсии, преобразуем его, найдём производную и стационарные точки у полученного выражения и проверим, что получили именно минимум.

1) Преобразуем дисперсию

$$\begin{aligned} M\{(x-\hat{x})^T(x-\hat{x})|y\} &= M\{(x-Ly)^T(x-Ly)|y = Cx+\omega\} \\ &= M\{(x-LCx-L\omega)^T(x-LCx-L\omega)\} \\ &= M\{x^T(I-LC)^T(I-LC)x - 2x^T(I-LC)^TL\omega + \omega^TL^TL\omega\} \\ &= M\left\{(x^T\ \omega^T)\begin{pmatrix}(I-LC)^T(I-LC) & -(I-LC)^TL \\ -L^T(I-LC) & L^TL\end{pmatrix}\begin{pmatrix}x \\ \omega\end{pmatrix}\right\} \end{aligned}$$

Воспользуемся свойством мат. ожидания квадратичной формы:

$$\begin{aligned} &(\bar{x}^T\ 0)\begin{pmatrix}(I-LC)^T(I-LC) & -(I-LC)^TL \\ -L^T(I-LC) & L^TL\end{pmatrix}\begin{pmatrix}\bar{x} \\ 0\end{pmatrix} \\ &\qquad + tr\left\{\begin{pmatrix}(I-LC)^T(I-LC) & -(I-LC)^TL \\ -L^T(I-LC) & L^TL\end{pmatrix}\begin{pmatrix}X & 0 \\ 0 & W\end{pmatrix}\right\} \\ &\qquad = \bar{x}^T(I-LC)^T(I-LC)\bar{x} \\ &\qquad + tr\left\{\begin{pmatrix}(I-LC)^T(I-LC)X & -(I-LC)^TLW \\ -L^T(I-LC)X & L^TLW\end{pmatrix}\right\} \\ &\qquad = tr\{(I-LC)\bar{x}\bar{x}^T(I-LC)^T\} + tr\{(I-LC)^T(I-LC)X\} \\ &\qquad + tr\{L^TLW\} = f(L) \end{aligned}$$

2) Ищем производную и приравниваем её к нулю

Распишем отдельно каждое слагаемое:

- Первое

$$tr\{\nabla((I-LC)\bar{x}\bar{x}^T(I-LC)^T)\} = tr\{-(I-LC)\bar{x}\bar{x}^TC^T - C\bar{x}\bar{x}^T(I-LC)^T\}$$
$$= -2tr\{C\bar{x}\bar{x}^T(I-LC)^T\}$$

- Второе

$$tr\{\nabla((I-LC)^T(I-LC))X\} = tr\{(-(I-LC)^TC - C^T(I-LC))X\}$$
$$= -2tr\{C^T(I-LC)X\} = -2tr\{(I-LC)XC^T\}$$
$$= -2tr\{CX(I-LC)^T\}$$

- Третье

$$tr\{\nabla(L^TL)W\} = tr\{(L^T+L)W\} = 2tr\{LW\} = 2tr\{WL^T\}$$

Объединяем и получаем следующее выражение:

$$\nabla f(L) = -2tr\{C\bar{x}\bar{x}^T(I-LC)^T\} - 2tr\{CX(I-LC)^T\} + 2tr\{LW\} = 0$$

- Пересоберём слагаемые

$$\nabla f(L) = -2tr\{C(\bar{x}\bar{x}^T + X)\} + 2tr\{(C\bar{x}\bar{x}^TC^T + CXC^T + W)L^T\} = 0$$

- Транспонируем выражения внутри обоих следов, чтобы размерность матрицы $L$ в дальнейшем была правильной

$$\nabla f(L) = -2tr\{(\bar{x}\bar{x}^T + X)C^T\} + 2tr\{L(C\bar{x}\bar{x}^TC^T + CXC^T + W)\} = 0$$

- Выразим $L$

$$L = (\bar{x}\bar{x}^T + X)C^T(C(\bar{x}\bar{x}^T + X)C^T + W)^{-1}$$

3) Проверяем, что получили именно минимум

$$\nabla^2 f(L) = W + C(X + \bar{x}\bar{x}^T)C^T > 0 \Rightarrow L - \text{т.}\min$$

Да уж, это оказалось сложнее, чем хотелось бы. Производная по матрице, использование свойств следа для преобразования, всё это нетривиальные вещи. Чтобы было понятнее, можете заглянуть в источники, а если вы сразу всё поняли, то вы потрясающий!

Что-то мы отвлеклись. Ну-ка взяли себя в руки, похлопали по щёчкам, переживать здесь больше не о чем, ведь мы уже всё сделали, осталось только записать полученную оценку:

$$\hat{x} = Ly = (X + \bar{x}\bar{x}^T)C^T(W + C(X + \bar{x}\bar{x}^T)C^T)^{-1}y$$

Мы нашли оценку состояния с помощью метода минимума дисперсии и учли поведение помехи, но эту формулу не назовёшь хорошей. При каждом следующем измерении нам нужно пересчитывать всё заново, как и в одношаговом методе минимума квадратов, и мы снова приходим к необходимости получения рекуррентной формулы.

Разберём один небольшой пример на дорожку.

Пример.

$$y_k = 2x + \omega_k, \qquad x \sim \mathcal{N}(0,1), \omega \sim \mathcal{N}(0,0.5)$$

Случайные величины на разных шагах не коррелированы.

$$y_1 = 2, \qquad y_2 = 1.8$$

Параметры:

$$X = 1, C = \begin{pmatrix} 2 & 2 \end{pmatrix}^T, W = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}, y = \begin{pmatrix} 2.0 & 1.8 \end{pmatrix}^T$$

Оценка будет иметь вид:

$$\hat{x} = \begin{pmatrix} 2 & 2 \end{pmatrix} \left( \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix} + \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix} \right)^{-1} \begin{pmatrix} 2.0 \\ 1.8 \end{pmatrix} = 0.894118$$

## Условное математическое ожидание и ковариационная матрица для гауссовских векторов

Мы хотим получить рекуррентную формулу для оценки состояния по методу минимума дисперсии. Как мы позднее увидим, для этой задачи нам нужно будет найти условное мат. ожидание и условную ковариационную матрицу для гауссовского вектора, поэтому разберёмся в этом сразу, чтобы потом быстро получить рекуррентную оценку.

Запишем постановку задачи:

$$y = Cx + \omega, \qquad \omega \sim \mathcal{N}(\overline{\omega}, W), \qquad x \sim \mathcal{N}(\bar{x}, X)$$

$$M\{(x - \bar{x})(\omega - \overline{\omega})^T\} = 0, x \in \mathbb{R}^n, \omega \in \mathbb{R}^m$$

$$f_{x|y}(x) = \frac{f_{xy}(x, y)}{f_y(y)}?$$

Чтобы задать условную плотность распределения $f_{x|y}(x)$ нужно найти условное мат. ожидание и условную ковариационную матрицу. Погнали.

Разобьём задачу на три этапа:

1. Найти плотность распределения $f_y(y)$
2. Найти совместную плотность распределения $f_{xy}(x, y)$
3. Собрать всё вместе и найти условную плотность распределения

Чтобы задать плотность нормального распределения нужно знать мат. ожидание, ковариационную матрицу, обратную к ней и определитель ковариационной матрицы. Поисками этих характеристик мы и займёмся.

1) Плотность распределения $f_y(y)$
   a) Мат. ожидание

$$\bar{y} = M\{y\} = C\bar{x} + \overline{\omega}$$

   b) Ковариационная матрица

$$y - \bar{y} = Cx + \omega - C\bar{x} - \overline{\omega} = C(x - \bar{x}) + (\omega - \overline{\omega}) = \begin{pmatrix} C & I \end{pmatrix} \begin{pmatrix} x - \bar{x} \\ \omega - \overline{\omega} \end{pmatrix}$$

$$Y = M\{(y - \bar{y})(y - \bar{y})^T\} = \begin{pmatrix} C & I \end{pmatrix} M\left\{ \begin{pmatrix} x - \bar{x} \\ \omega - \overline{\omega} \end{pmatrix} \begin{pmatrix} (x - \bar{x})^T & (\omega - \overline{\omega})^T \end{pmatrix} \right\} \begin{pmatrix} C^T \\ I \end{pmatrix}$$
$$= \begin{pmatrix} C & I \end{pmatrix} \begin{pmatrix} X & 0 \\ 0 & W \end{pmatrix} \begin{pmatrix} C^T \\ I \end{pmatrix} = CXC^T + W$$

   c) Обратная к ковариационной матрице:

$$Y^{-1} = (CXC^T + W)^{-1}$$

Используем тождество Шермана-Моррисона-Вудбери, чтобы преобразовать данное выражение. Подробнее о нём можно узнать в приложении.

$$Y^{-1} = W^{-1} - W^{-1}C(X^{-1} + C^TW^{-1}C)^{-1}C^TW^{-1}$$

d) Определитель ковариационной матрицы

$$\det Y = \det(CXC^T + W)$$

2) Совместная плотность распределения

Для удобной работы введём новый вектор $z = (x^T\ y^T)^T$

a) Мат. ожидание

$$\bar{z} = M\{z\} = \begin{pmatrix} \bar{x} \\ C\bar{x} + \bar{\omega} \end{pmatrix}$$

b) Ковариационная матрица

$$z - \bar{z} = \begin{pmatrix} x - \bar{x} \\ C(x - \bar{x}) + (\omega - \bar{\omega}) \end{pmatrix} = \begin{pmatrix} I & 0 \\ C & I \end{pmatrix} \begin{pmatrix} x - \bar{x} \\ \omega - \bar{\omega} \end{pmatrix}$$

$$\begin{aligned} Z &= M\{(z - \bar{z})(z - \bar{z})^T\} \\ &= \begin{pmatrix} I & 0 \\ C & I \end{pmatrix} M\left\{\begin{pmatrix} x - \bar{x} \\ \omega - \bar{\omega} \end{pmatrix} \begin{pmatrix} (x - \bar{x})^T & (\omega - \bar{\omega})^T \end{pmatrix}\right\} \begin{pmatrix} I & C^T \\ 0 & I \end{pmatrix} \\ &= \begin{pmatrix} I & 0 \\ C & I \end{pmatrix} \begin{pmatrix} X & 0 \\ 0 & W \end{pmatrix} \begin{pmatrix} I & C^T \\ 0 & I \end{pmatrix} = \begin{pmatrix} X & 0 \\ CX & W \end{pmatrix} \begin{pmatrix} I & C^T \\ 0 & I \end{pmatrix} \\ &= \begin{pmatrix} X & XC^T \\ CX & CXC^T + W \end{pmatrix} \end{aligned}$$

c) Обратная к ковариационной матрице

$$\begin{aligned} Z^{-1} &= \left(\begin{pmatrix} I & 0 \\ C & I \end{pmatrix} \begin{pmatrix} X & 0 \\ 0 & W \end{pmatrix} \begin{pmatrix} I & C^T \\ 0 & I \end{pmatrix}\right)^{-1} = \begin{pmatrix} I & -C^T \\ 0 & I \end{pmatrix} \begin{pmatrix} X^{-1} & 0 \\ 0 & W^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -C & I \end{pmatrix} \\ &= \begin{pmatrix} X^{-1} & -C^TW^{-1} \\ 0 & W^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -C & I \end{pmatrix} = \begin{pmatrix} X^{-1} + C^TW^{-1}C & -C^TW^{-1} \\ -W^{-1}C & W^{-1} \end{pmatrix} \end{aligned}$$

d) Определитель ковариационной матрицы

Определитель этой матрицы найти не так уж просто. Для этого придётся сделать преобразование Шура и привести матрицу к блочно-диагональному виду.

$$\det Z = \det U^TZU, \qquad \det U = 1$$

Выберем матрицу следующим образом:

$$U = \begin{pmatrix} I & 0 \\ \alpha & I \end{pmatrix}$$

Это сделано для того, чтобы правый нижний блок остался не тронутым, что впоследствии поможет упростить выражение

$$U^TZU = \begin{pmatrix} I & \alpha^T \\ 0 & I \end{pmatrix} \begin{pmatrix} X & XC^T \\ CX & CXC^T + W \end{pmatrix} \begin{pmatrix} I & 0 \\ \alpha & I \end{pmatrix}$$
$$= \begin{pmatrix} X + XC^T\alpha + \alpha^T CX + \alpha^T(CXC^T + W)\alpha & XC^T + \alpha^T(CXC^T + W) \\ CX + (CXC^T + W)\alpha & CXC^T + W \end{pmatrix}$$

Выбираем $\alpha$ такое, чтобы элементы на побочной диагонали занулились:

$$CX + (CXC^T + W)\alpha = 0 \Rightarrow \alpha = -(CXC^T + W)^{-1}CX$$

$$\alpha^T = -XC^T(CXC^T + W)^{-1}$$

$$\begin{aligned} XC^T\alpha + \alpha^T CX + \alpha^T(CXC^T + W)\alpha \\ &= -2XC^T(CXC^T + W)^{-1}CX \\ &+ XC^T(CXC^T + W)^{-1}(CXC^T + W)(CXC^T + W)^{-1}CX \\ &= -XC^T(CXC^T + W)^{-1}CX \end{aligned}$$

Запишем полученную матрицу:

$$U^TZU = \begin{pmatrix} X - XC^T(CXC^T + W)^{-1}CX & 0 \\ 0 & CXC^T + W \end{pmatrix}$$
$$= \begin{pmatrix} (X^{-1} + C^TW^{-1}C)^{-1} & 0 \\ 0 & CXC^T + W \end{pmatrix}$$

Осталось только посчитать записать определитель:

$$\det Z = \det U^TZU = \det(X^{-1} + C^TW^{-1}C)^{-1}\det(CXC^T + W)$$

3) Условная плотность распределения

$$f_{xy}(x) = \frac{1}{(2\pi)^{\frac{n+m}{2}}\sqrt{\det \mathrm{Z}}} e^{-\frac{1}{2}(z-\bar{z})^T\mathrm{Z}^{-1}(z-\bar{z})}$$

$$f_y(y) = \frac{1}{(2\pi)^{\frac{m}{2}}\sqrt{\det \mathrm{Y}}} e^{-\frac{1}{2}(y-\bar{y})^T\mathrm{Y}^{-1}(y-\bar{y})}$$

Подставляем, отступать некуда.

$$f_{x|y}(x) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det \mathrm{X_y}}} e^{-\frac{1}{2}(x-\bar{x}_y)^T\mathrm{X}_\mathrm{y}^{-1}(x-\bar{x}_y)} = \frac{(2\pi)^{\frac{m}{2}}\sqrt{\det \mathrm{Y}}}{(2\pi)^{\frac{n+m}{2}}\sqrt{\det \mathrm{Z}}} e^{-\frac{1}{2}(\dots)}$$

$$(\dots) = (z-\bar{z})^T\mathrm{Z}^{-1}(z-\bar{z}) - (y-\bar{y})^T\mathrm{Y}^{-1}(y-\bar{y})$$

a) Преобразуем выражение перед экспонентой

$$\frac{(2\pi)^{\frac{m}{2}}\sqrt{\det \mathrm{Y}}}{(2\pi)^{\frac{n+m}{2}}\sqrt{\det \mathrm{Z}}} = \frac{\sqrt{\det(CXC^T+W)}}{(2\pi)^{\frac{n}{2}}\sqrt{\det(X^{-1}+C^TW^{-1}C)^{-1}\det(CXC^T+W)}}$$
$$= \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det(X^{-1}+C^TW^{-1}C)^{-1}}}$$

Отсюда мы сразу получаем условную ковариационную матрицу:

$$X_y = (X^{-1}+C^TW^{-1}C)^{-1}$$

На самом деле, мы не можем из равенства определителей так просто заключить, что это именно ковариационная матрица, поэтому мы проверим это, когда будем преобразовывать степень экспоненты.

b) Преобразуем степень экспоненты
   - Первое слагаемое

$$Z^{-1} = \begin{pmatrix} X^{-1}+C^TW^{-1}C & -C^TW^{-1} \\ -W^{-1}C & W^{-1} \end{pmatrix}$$

$$\begin{aligned}&(z-\bar{z})^T\mathrm{Z}^{-1}(z-\bar{z})\\ &\quad= \begin{pmatrix}(x-\bar{x})^T & (y-\bar{y})^T\end{pmatrix}\begin{pmatrix} X^{-1}+C^TW^{-1}C & -C^TW^{-1} \\ -W^{-1}C & W^{-1} \end{pmatrix}\begin{pmatrix} x-\bar{x} \\ y-\bar{y}\end{pmatrix}\\ &\quad= (x-\bar{x})^T(X^{-1}+C^TW^{-1}C)(x-\bar{x}) - (x-\bar{x})^TC^TW^{-1}(y-\bar{y})\\ &\quad- (y-\bar{y})^TW^{-1}C(x-\bar{x}) + (y-\bar{y})^TW^{-1}(y-\bar{y})\end{aligned}$$

   - Второе слагаемое

$$Y^{-1} = W^{-1} - W^{-1}C(X^{-1}+C^TW^{-1}C)^{-1}C^TW^{-1}$$

$$\begin{aligned}&(y-\bar{y})^T\mathrm{Y}^{-1}(y-\bar{y})\\ &\quad= (y-\bar{y})^T\mathrm{W}^{-1}(y-\bar{y})\\ &\quad- (y-\bar{y})^TW^{-1}C(X^{-1}+C^TW^{-1}C)^{-1}C^TW^{-1}(y-\bar{y})\end{aligned}$$

   - Берём разность

Воспользуемся нашим предположением о виде условной ковариационной матрицы и перепишем выражения с ней

$$\begin{aligned}&(z-\bar{z})^T\mathrm{Z}^{-1}(z-\bar{z}) - (y-\bar{y})^T\mathrm{Y}^{-1}(y-\bar{y})\\ &\quad= (x-\bar{x})^TX_y^{-1}(x-\bar{x}) - (x-\bar{x})^TC^TW^{-1}(y-\bar{y})\\ &\quad- (y-\bar{y})^TW^{-1}C(x-\bar{x}) + (y-\bar{y})^TW^{-1}CX_yC^TW^{-1}(y-\bar{y})\end{aligned}$$

   - Сгруппируем слагаемые

$$(x-\bar{x})^T X_y^{-1}\left((x-\bar{x})-X_y C^T W^{-1}(y-\bar{y})\right) - (y-\bar{y})^T W^{-1} C X_y X_y^{-1}\left((x-\bar{x})-X_y C^T W^{-1}(y-\bar{y})\right)$$

- Вынесем общий множитель

$$\left((x-\bar{x})^T-(y-\bar{y})^T W^{-1} C X_y\right) X_y^{-1}\left((x-\bar{x})-X_y C^T W^{-1}(y-\bar{y})\right) = \left(x-\bar{x}_y\right)^T \mathrm{X}_y^{-1}\left(x-\bar{x}_y\right)$$

Вот всё и совпало, наша ковариационная матрица стоит там, где и должна, а значит наше предположение было верным. Осталось выразить отсюда условное мат. ожидание.

$$x-\bar{x}-X_y C^T W^{-1}(y-\bar{y}) = x-\bar{x}_y \Rightarrow \bar{x}_y=\bar{x}+X_y C^T W^{-1}(y-\bar{y})$$

Подставим $\bar{y}$ и получим окончательный ответ:

$$\bar{x}_y=\bar{x}+X_y C^T W^{-1}(y-C\bar{x}-\bar{\omega}), \qquad X_y=(X^{-1}+C^T W^{-1} C)^{-1}$$

Вот мы и нашли условное мат. ожидание и условную ковариационную матрицу. Было почти не сложно. Двигаемся дальше, впереди самое интересное.

## Рекуррентное оценивание по методу минимума дисперсии

Как мы помним, нам нужно получить рекуррентную формулу для оценки состояния системы с учётом помехи и для этого мы хотим использовать метод минимума дисперсии.

Итак, формальный вид задачи следующий:

$$y_k = Cx + \omega_k, \qquad x \in \mathbb{R}^n;\ y, \omega \in \mathbb{R}^m$$

$$\omega_k \sim \mathcal{N}(\overline{\omega}_k, W_k), \qquad M\left\{(\omega_k - \overline{\omega}_k)\left(\omega_j - \overline{\omega}_j\right)^T\right\} = \begin{cases} 0, k \neq j \\ W_k, k = j \end{cases}$$

$$x \sim \mathcal{N}(\bar{x}, X), \qquad M\{(\omega_k - \overline{\omega}_k)(x - \bar{x})^T\} = 0$$

Нужно найти рекуррентную оценку по методу минимума дисперсии:

$$\hat{x}_k = \arg\min_{\hat{x}} M\{(x - \hat{x})^T (x - \hat{x}) | y_1, \dots, y_k\}$$

Если присмотреться к формуле, то становится ясно, что аргументом этого минимума будет условное мат. ожидание:

$$\hat{x}_k = M\{x | y_1, \dots, y_k\}$$

Ясно, да? Вот это действительно строгое доказательство, здорово. Не переживайте, сейчас всё докажем.

Доказательство:

1) Запишем выражение в виде функции

$$\arg\min_{\hat{x}} M\{(x - \hat{x})^T (x - \hat{x}) | y_1, \dots, y_k\} = \min_{\hat{x}} f(\hat{x})$$

2) Найдём стационарную точку

$$\nabla f(\hat{x}) = -2M\{(x - \hat{x}) | y_1, \dots, y_k\} = 0 \Rightarrow \hat{x}^* = M\{x | y_1, \dots, y_k\}$$

3) Проверим, что нашли именно минимум

$$\nabla^2 f(\hat{x}) = 2I > 0 \Rightarrow \hat{x}^* - \text{т} \min.$$

Отлично, мы нашли выражение для оценки на $k$-ом шаге, осталось задать рекуррентную формулу. Для этого введём новую случайную величину, точнее сказать новые:

$$\breve{x}_k \sim \mathcal{N}(\hat{x}_k, X_k)$$

$$\hat{x}_k = M\{x | y_1, \dots, y_k\}, \qquad X_k = M\{(\breve{x}_k - \hat{x}_k)(\breve{x}_k - \hat{x}_k)^T\}$$

Теперь выразим оценку на $k + 1$-ом шаге, используя эту случайную величину:

$$\hat{x}_{k+1} = M\{x | y_1, \dots, y_{k+1}\} = M\{\breve{x}_k | y_{k+1}\}$$

Это уже известное нам условное мат. ожидание:

$$M\{x|y\} = \bar{x} + X_y C^T W^{-1}(y - C\bar{x} - \overline{\omega})$$

$$X_y = (X^{-1} + C^T W^{-1} C)^{-1} = M\left\{\left(x - \bar{x}_y\right)\left(x - \bar{x}_y\right)^T\right\}$$

И это значит, что мы можем записать рекуррентную формулу:

$$\hat{x}_{k+1} = \hat{x}_k + X_{k+1} C^T W_{k+1}^{-1}(y_{k+1} - C\hat{x}_k - \overline{\omega}_{k+1})$$

$$X_{k+1} = (X_k^{-1} + C^T W_{k+1}^{-1} C)^{-1}$$

*Важно.* Исходя из построения оценки следует, что $\hat{x}_0 = \bar{x}, X_0 = X$

Вот и всё. С предварительной подготовкой эта задача оказалась довольно простой. Посмотрим, что же будет дальше.

## Постановка задачи оптимальной фильтрации. Фильтр Калмана

Итак, мы с вами научились оценивать состояние системы по выходу, пора бы уже и применить это на практике.

Допустим нам задано уравнение динамики системы и задан выход:

$$\begin{cases} x_{k+1} = Ax_k + v_k \\ y_k = Cx_k + \omega_k \end{cases}$$

$$x_0 \sim \mathcal{N}(\bar{x}_0, X_0), v_k \sim \mathcal{N}(\bar{v}_k, V_k), \omega_k \sim \mathcal{N}(\bar{\omega}_k, W_k)$$

Случайные величины не коррелированы между собой.

Как мы видим, в обоих уравнениях есть помехи, которые мешают правильно оценить текущее состояние. Нам хочется от этого избавиться и получать чистое состояния, без помех, насколько это возможно. Более формально это можно представить так:

$$\hat{x}_k = \arg\min_{\hat{x}} M\{(x_k - \hat{x})^T (x_k - \hat{x}) | y_1, \dots, y_k\}$$

Такая оценка реализуется специальным наблюдателем – фильтром Калмана.

*Определение.* $\hat{x}_k^-$ - априорная оценка состояния $x_k$, полученная без учёта результата измерения $k$-ого шага

$$\hat{x}_k^- = M\{x_k | y_1, \dots, y_{k-1}\}$$

*Определение.* $\hat{x}_k^+$ - апостериорная оценка состояния $x_k$, полученная по результатам измерения на $k$-ом шаге.

$$\hat{x}_k^+ = M\{x_k | y_1, \dots, y_k\}$$

Формулы выводятся аналогично тому, как мы это делали в предыдущем разделе.

Установим связь между этими оценками, для чего поставим в выражение для априорной оценки уравнение динамики:

$$\hat{x}_{k+1}^- = M\{x_{k+1} | y_1, \dots, y_k\} = M\{Ax_k + v_k | y_1, \dots, y_k\} = A\hat{x}_k^+ + \bar{v}_k$$

Введём ещё два понятия.

*Определение.* $E_k^-$ - ковариационная матрица априорной ошибки на $k$-ом шаге

$$E_k^- = M\{(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T\}$$

*Определение.* $E_k^+$ - ковариационная матрица апостериорной ошибки на $k$-ом шаге

$$E_k^+ = M\{(x_k - \hat{x}_k^+)(x_k - \hat{x}_k^+)^T\}$$

Установим также связь между этими формулами.

$$\begin{aligned} E_{k+1}^{-} &= M\{(x_{k+1} - \hat{x}_{k+1}^{-})(x_{k+1} - \hat{x}_{k+1}^{-})^T\} \\ &= M\{(A(x_k - \hat{x}_k^{+}) + v_k - \bar{v}_k)(A(x_k - \hat{x}_k^{+}) + v_k - \bar{v}_k)^T\} \\ &= AM\{(x_k - \hat{x}_k^{+})(x_k - \hat{x}_k^{+})^T\}A^T + M\{(v_k - \bar{v}_k)(v_k - \bar{v}_k)^T\} \\ &= AE_k^{+}A^T + V_k \end{aligned}$$

Отлично, осталось замкнуть круг и выразить апостериорные оценки через априорные. Тут всё тоже просто, сделаем тоже самое, что и при выводе рекуррентной формулы по методы минимума дисперсии. Введём новую случайную величину:

$$x_k^{-} \sim \mathcal{N}(\hat{x}_k^{-}, E_k^{-})$$

Тогда можем записать апостериорную оценку через априорную:

$$\hat{x}_k^{+} = M\{x_k | y_1, \dots, y_k\} = M\{x_k^{-} | y_k\}$$

И сразу найти выражения для апостериорных оценок:

$$\hat{x}_k^{+} = \hat{x}_k^{-} + E_k^{+}C^T W_k^{-1}(y_k - C\hat{x}_k^{-} - \bar{\omega}_k)$$

$$E_k^{+} = ((E_k^{-})^{-1} + C^T W_k^{-1} C)^{-1}$$

Объединим всё, что нам удалось получить:

$$\begin{cases} E_k^{-} = AE_{k-1}^{+}A^T + V_{k-1} \\ E_k^{+} = ((E_k^{-})^{-1} + C^T W_k^{-1} C)^{-1} \\ \hat{x}_k^{-} = A\hat{x}_{k-1}^{+} + \bar{v}_{k-1} \\ \hat{x}_k^{+} = \hat{x}_k^{-} + E_k^{+}C^T W_k^{-1}(y_k - C\hat{x}_k^{-} - \bar{\omega}_k) \end{cases}$$

*Важно.* Исходя из построения, можно сказать, что $\hat{x}_0^{+} = \bar{x}_0, E_0^{+} = X_0$

Это и есть фильтр Калмана. Ура, товарищи! Ура!

Напоследок введём одно полезное определение, которое может часто встречаться в литературе.

*Определение.* Матрица $K_k = E_k^{+}C^T W_k^{-1}$ называется коэффициентом усиления фильтра Калмана.

## Стохастическая оптимальная линейная система при неполной информации о состоянии. Задача линейно-квадратичного гауссовского управления. Принцип разделения. Общая структура решения

Мы уже рассматривали задачу стохастического оптимального управления, но при полной информации о состоянии, сейчас мы усложним задачу и получим решение при неполной информации о состоянии.

Итак, запишем задачу:

$$\begin{cases} x_{k+1} = Ax_k + Bu_k + v_k \\ y_k = Cx_k + \omega_k \end{cases}, k = 0,1, \dots, N-1$$

$$x_0 \sim \mathcal{N}(\bar{x}_0, X_0), v_k \sim \mathcal{N}(\bar{v}_k, V_k), \omega_k \sim \mathcal{N}(\bar{\omega}_k, W_k)$$

Случайные величины не коррелированы между собой.

Нужно минимизировать функционал

$$J = M\left\{x_N^T S x_N + \sum_{i=0}^{N-1} \left(x_i^T Q x_i + u_i^T R u_i\right)\right\} \to \min_{u_0,\dots,u_{N-1}}$$

$$S = S^T \geq 0, \qquad Q = Q^T \geq 0\,, \qquad R = R^T > 0$$

Решение такой задачи можно разделить на два отдельных этапа, сущность этого разделения мы представим в виде теоремы.

**Теорема о разделении**. Решение задачи линейно-квадратичного гауссовского управления при неполной информации о состоянии может быть получено в два этапа

1) Ищем решение задачи линейно-квадратичного управления

$$u_k^* = -(R + B^T X_{k+1} B)^{-1} B^T X_{k+1} A x_k$$

2) Строим фильтр Калмана и используем полученные им оценки для замыкания управления

$$u_k^* = -(R + B^T X_{k+1} B)^{-1} B^T X_{k+1} A \hat{x}_k^+ = -M \hat{x}_k^+$$

Без доказательства.

С таким управлением система примет вид:

$$x_{k+1} = Ax_k - BM\hat{x}_k^+ + v_k, \qquad k = 0,1, \dots, N-1$$

$$\hat{x}_k^+ = \hat{x}_k^- + E_k^+ C^T W_k^{-1}(y_k - C\hat{x}_k^- - \bar{\omega}_k), \qquad E_k^+ = ((E_k^-)^{-1} + C^T W_k^{-1} C)^{-1}$$

Обновим априорные оценки для данной задачи, так как уравнение динамики изменилось.

a) Априорная оценка состояния

$$\begin{aligned}\hat{x}_{k+1}^{-} &= M\{x_{k+1}|y_1, \dots, y_k\} = M\{Ax_k - BM\hat{x}_k^{+} + v_k|y_1, \dots, y_k\} \\ &= (A - BM)\hat{x}_k^{+} + \bar{v}_k\end{aligned}$$

b) Ковариационная матрица априорной ошибки

$$\begin{aligned}x_{k+1} - \hat{x}_{k+1}^{-} &= Ax_k - BM\hat{x}_k^{+} + v_k - (A - BM)\hat{x}_k^{+} - \bar{v}_k \\ &= A(x_k - \hat{x}_k^{+}) + v_k - \bar{v}_k\end{aligned}$$

$$\begin{aligned}E_{k+1}^{-} = M\{(x_{k+1} - \hat{x}_{k+1}^{-})(x_{k+1} - \hat{x}_{k+1}^{-})^T\} \\ &= M\{(A(x_k - \hat{x}_k^{+}) + v_k - \bar{v}_k)(A(x_k - \hat{x}_k^{+}) + v_k - \bar{v}_k)^T\} \\ &= AM\{(x_k - \hat{x}_k^{+})(x_k - \hat{x}_k^{+})^T\}A^T + M\{(v_k - \bar{v}_k)(v_k - \bar{v}_k)^T\} \\ &= AE_k^{+}A^T + V_k\end{aligned}$$

Да, ковариационная матрица осталась прежней.

Как было сказано ранее, теорема приводится без доказательства, но мы всё же посмотрим саму идею, чтобы понимать законность наших действий.

Введём новую величину – погрешность оценки:

$$e_k = x_k - \hat{x}_k^{+}$$

Это нормально распределённая случайная величина:

$$e_k \sim \mathcal{N}(0, E_k^{+})$$

$$M\{e_k\} = M\{x_k - \hat{x}_k^{+}|y_1, \dots, y_k\} = 0$$

$$M\{e_k e_k^T\} = M\{(x_k - \hat{x}_k^{+})(x_k - \hat{x}_k^{+})^T\} = E_k^{+}$$

Получается, если $E_k^{+}$ уменьшается с каждым шагом, то и ошибка уменьшается, а значит оценочные значения сходятся к реальному состоянию системы. И да, $E_k^{+}$ действительно уменьшается с каждым шагом, но доказывать мы этого не будем.

^_^

## Линейно-матричные неравенства. Связь с уравнением Ляпунова

Сколько интересного мы с вами успели пройти, это было действительно здорово. Но не будем останавливаться на достигнутом, продолжим усердно трудится и освоим ещё одну важную тему – линейно-матричные неравенства.

*Определение.* Линейно-матричным неравенством ($LMI$) называется выражение вида

$$F(x) < (>)\ 0, \qquad F = F^T \in \mathbb{R}^{m\times m}$$

В дальнейшем мы всё в этой главе будем записывать для знака меньше, но если заменить знак на больше, то смысл не изменится.

*Примечание.*

1) Матрица $F$ симметрична, следовательно все её собственные числа действительны.
2) Знак меньше (больше) здесь означает знакоопределённость матрицы.

Линейно-матричные неравенства имеют две основные формы записи.

a. Каноническая форма записи

$$F(x) = F_0 + \sum_{k=1}^{n} F_k x_k < \ 0, \qquad x \in \mathbb{R}^n, F_k = F_k^T \in \mathbb{R}^{m\times m}, k = 0, \dots, n$$

b. Стандартная форма записи

$$F(X) < 0, \qquad X = X^T \in \mathbb{R}^{n\times n}\ (n \le m)$$

Пример

$$A^T X + XA < 0$$

Стандартную форму записи можно свести к канонической с помощью разложения по базису:

$$X = \sum_{k=1}^{n(n+1)/2} E_k x_k$$

$$E_1 = \begin{pmatrix} 1 & 0 & \dots \\ 0 & \dots & \dots \\ \dots & \dots & \dots \end{pmatrix}, E_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & \dots \\ 0 & \dots & \dots \end{pmatrix}, E_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & \dots \\ 1 & \dots & \dots \end{pmatrix}, \dots$$

Линейно-матричные неравенства обладают рядом свойств, которые нужно хорошенько запомнить. Давайте постараемся.

Итак, свойства.

1) $D = \{x \in \mathbb{R}^n : F(x) < 0\}$ – выпуклая область

Доказательство:

Множество $D$ выпуклое $\Leftrightarrow$ $\forall \alpha \in [0; 1]\ \forall x, y \in D: \alpha x + (1-\alpha)y \in D$

$$\begin{aligned} F(\alpha x + (1-\alpha)y) &= (\alpha + 1 - \alpha)F_0 + \sum_{k=1}^{n} F_k(\alpha x_k + (1-\alpha)y_k) \\ &= \alpha\left(F_0 + \sum_{k=1}^{n} F_k x_k\right) + (1-\alpha)\left(F_0 + \sum_{k=1}^{n} F_k y_k\right) \\ &= \alpha F(x) + (1-\alpha)F(y) < 0 \Rightarrow\ \alpha x + (1-\alpha)y \in D \end{aligned}$$

2) Объединение системы $LMI$ в одно

$$G_k(x) < 0, k = 1, \dots m \Rightarrow F(x) = diag(G_1, \dots, G_m) < 0$$

3) Обращение знака

$$F(x) < 0 \Leftrightarrow -F(x) > 0$$

4) Сохранение знака при преобразовании

$$\forall S \in \mathbb{R}^{m\times m}: \det S \neq 0, F(x) < 0 \Rightarrow S^T F(x) S < 0, \qquad S^{-1}F(x)S < 0$$

5) Связь знакоопределённости матрицы и собственных чисел

$$F(x) < 0 \Leftrightarrow \lambda_{max}(F) < 0$$

6) Оценка квадратичной формы

$$\lambda_{min}(A)x^T x \leq x^T A x \leq \lambda_{max}(A)x^T x$$

Доказательство:

$$x = Sy, S^T S = I \Rightarrow y^T S^T A S y = y^T D y, D = diag\big(\lambda_1(A), \dots, \lambda_n(A)\big)$$

$$\lambda_{min}(A)y^T y \leq y^T D y = \sum_{k=1}^{n} y_k^2 \lambda_k \leq \lambda_{max}(A)y^T y \Leftrightarrow$$

$$\Leftrightarrow \lambda_{min}(A)x^T x \leq x^T A x \leq \lambda_{max}(A)x^T x$$

Да, доказали мы только первое и шестое свойства, но остальные утверждения очевидны.

Ну что же, с основными понятиями и свойствами мы познакомились, пора бы и посмотреть, зачем нам всё это надо. Для этого рассмотрим несколько примеров конкретных задач.

a) Отыскание максимального собственного числа заданной матрицы

$$\lambda_{max}(F) = \arg\inf_{\gamma}(F - \gamma I \leq 0)$$

То есть, используя линейно-матричные неравенства, мы можем найти максимальное собственное число матрицы. Да, есть варианты и лучше, но сам факт того, что это возможно, вполне интересен.

Давайте теперь посмотрим, как можно применить $LMI$ к уже известным нам задачам из курса. Чтобы понимать последующие примеры, вспомним одну важную теорему.

**Теорема.** Решение $x = 0$ системы $\dot{x} = Ax$ является асимптотически устойчивым по Ляпунову тогда и только тогда, когда

$$\exists X = X^T > 0: \quad A^T X + XA < 0$$

Ну а теперь продолжим. Ещё одним примером использования задач оптимизации являются задачи синтеза стабилизирующего управления.

b) Линейная система с полной информацией о состоянии

Дана система

$$\dot{x} = Ax + Bu$$

мы хотим стабилизировать её, используя управление в виде линейной обратной связи по состоянию: $u = \theta x$. Матрица замкнутой системы будет иметь вид:

$$A_c = A + B\theta$$

Найти такое управление можно с помощью линейно-матричных неравенств:

$$A_c = A + B\theta - \text{асимпт. уст.} \Leftrightarrow \exists X = X^T > 0: A_c^T X + XA_c < 0$$

- Преобразуем это выражение

$$A_c^T X + XA_c < 0 \Leftrightarrow A^T X + \theta^T B^T X + XA + XB\theta < 0$$

Как мы видим, оно на самом деле нелинейное, потому что нам неизвестно $X$ и $\theta$ и есть слагаемые, в которых они встречаются вместе. Чтобы всё-таки смести это неравенство к $LMI$ помножим выражение справа и слева на $X^{-1}$:

$$X^{-1}A^T + X^{-1}\theta^T B^T + AX^{-1} + B\theta X^{-1} < 0$$

- Сделаем замену

$$Y = X^{-1}, \qquad Z = \theta Y$$

Тогда наше выражение станет линейным:

$$YA^T + Z^T B^T + AY + BZ < 0$$

А матрицу $\theta$ можно легко найти:

$$\theta = ZY^{-1}$$

c) Линейная система с неполной информацией о состоянии

Задана система

$$\begin{cases} \dot{x} = Ax + Bu \\ \quad y = Cx \end{cases}$$

Мы не знаем полного состояния системы, известен только выход - $y$. Для решения этой задачи мы построим линейный асимптотический наблюдатель:

$$\dot{\xi} = A\xi + Bu + LC(x - \xi)$$

а управление будем искать в виде линейной обратной связи по состоянию наблюдателя: $u = \theta\xi$.

Из предыдущих разделов курса нам известно, что для сходимости наблюдателя к состоянию системы и асимптотической устойчивости системы должно выполняться два условия:

- $A + B\theta$ – асимптотически устойчива
- $A - LC$ – асимптотически устойчива

И найти неизвестные матрицы можно с помощью линейно-матричных неравенств:

1) $(A + B\theta)^T X + X(A + B\theta) < 0$

Такое неравенство мы рассматривали выше и получили решение:

$$YA^T + Z^T B^T + AY + BZ < 0, \qquad Y = X^{-1}, Z = \theta Y$$

$$\theta = ZY^{-1}$$

2) $(A - LC)^T X + X(A - LC) < 0$

Здесь всё аналогично:

$$A^T X - C^T L^T X + XA - XLC < 0$$

- Сделаем замену

$$W = XL$$

И выражение преобразуется к нужному нам $LMI$

$$A^T X - C^T W^T + XA - WC < 0$$

А неизвестная матрица $L$ легко выражается из применённой замены:

$$L = X^{-1} W$$

На этом мы заканчиваем данную тему, но не заканчиваем рассмотрение линейно-матричных неравенств. В следующей главе мы введём несколько новых определений, одним из которых будет новое понятие устойчивости, так что не расходимся, много интересного впереди.

## $LMI$-область: определение и свойства. Понятие $D$-устойчивости линейной системы. Теорема об устойчивости относительно $LMI$-области

Сейчас Вы узнаете, что же такое $LMI$-область, что такое произведение Кронекера и как можно сдвигать собственные числа в те области, которые хотите, ну почти. Поехали.

*Определение.* Множество $D = \{z \in \mathbb{C}: F(z) < (>)\ 0\}$ называется $LMI$-областью, если

$$\exists L = L^T, M \in \mathbb{R}^{m \times m}:\ F(z) = L + Mz + M^T \bar{z} < (>)\ 0$$

При это $F(z)$ называют характеристической функцией.

$LMI$-область обладает парочкой важных свойств:

1) Симметрия относительно действительной оси

Доказательство:

$$F(z) = F^T(\bar{z})$$

2) Объединение нескольких областей в одну

$$D = \bigcap_{i=1}^{n} D_i, \qquad \forall i = 1, \dots, n: D_i = \{z \in \mathbb{C}: F_i(z) = L_i + M_i z + M_i^T \bar{z} < (>)\ 0\}$$

Тогда $D = \{z \in \mathbb{C}: F(z) = L + Mz + M^T \bar{z} < (>)\ 0\}$ – $LMI$-область, где

$$L = diag(L_1, L_2, \dots, L_n), M = diag(M_1, M_2, \dots, M_n)$$

Теперь можем поговорить о том, чтобы двигать собственные числа. Как вы, наверно, догадываетесь, двигать собственные числа мы можем не куда угодно, а только в $LMI$-области. Да, это достаточно серьёзное ограничение, но как мы чуть позже увидим, этого нам хватит с лихвой.

*Определение.* Матрица $A \in \mathbb{R}^{m \times m}$ является $D$-устойчивой, если все её собственные числа принадлежат области $D$:

$$spec\ A \subseteq int\ D$$

где $spec\ A$ – собственные числа матрицы $A$, $int\ D$ – внутренность области $D$

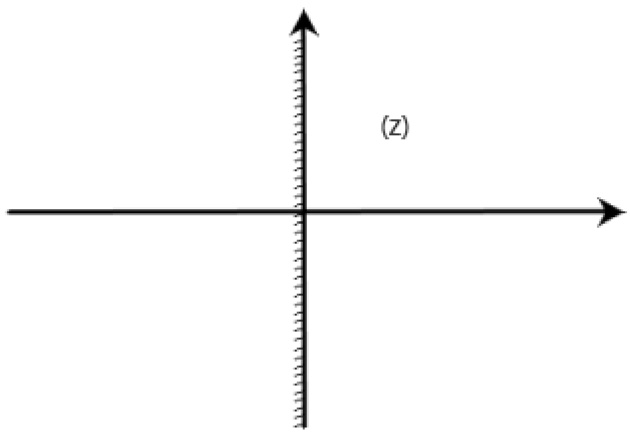
$D$-устойчивые матрицы относительно левой полуплоскости и единичного круга даже имеют свои названия.

*Определение.* Матрица $A$: $spec\ A \in \{Rez\} < 0$ называется гурвицевой

*Определение.* Матрица $A$: $spec\ A \in \{|z| < 1\}$ называется шуровской

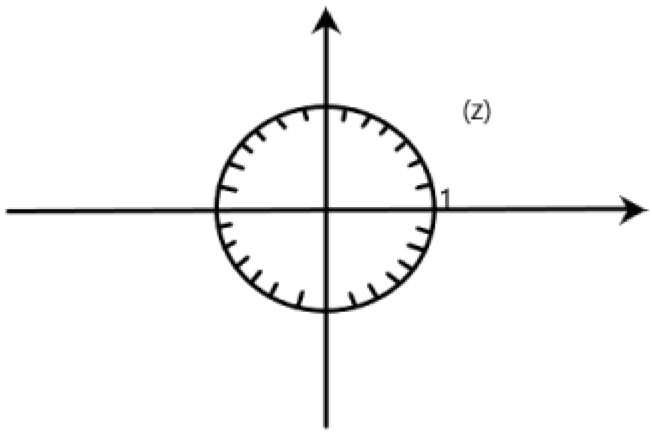
Теперь давайте рассмотрим несколько примеров $LMI$-областей.

1) Левая полуплоскость: $D = \{z \in \mathbb{C}: Rez < 0\}$




$$Rez = \frac{z + \bar{z}}{2} < 0 \Leftrightarrow z + \bar{z} < 0 \Rightarrow \ L = 0, M = 1$$

2) Внутренность единичного круга: $D = \{z \in \mathbb{C}: |z| < 1\}$




$$|z| < 1 \Leftrightarrow |z|^2 < 1 \Leftrightarrow 1 - z\bar{z} > 0$$

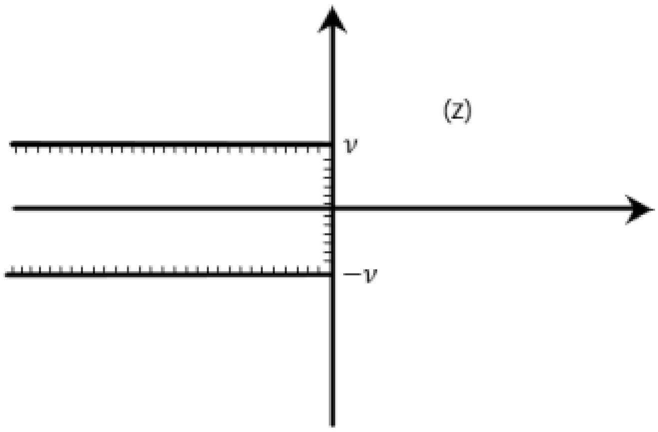
Чтобы свести это выражение к $LMI$-области используем критерий Сильвестра. Для этого нужно представить это выражение как определитель матрицы, ну или один из её миноров. В данном случае

$$\begin{pmatrix} 1 & z \\ \bar{z} & 1 \end{pmatrix} > 0 \Leftrightarrow 1 > 0, 1 - z\bar{z} > 0$$

$$L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Если кто-то забыл, что такое критерий Сильвестра и как он работает, то вам дорога в Интернет, можно перейти по ссылке в источниках.

3) Горизонтальная полуполоса: $D = \{z \in \mathbb{C}: Rez < 0, |Imz| < v\}$




a) $|Imz| < v$

$$|Imz| < v \Leftrightarrow \left|\frac{z-\bar{z}}{2i}\right|^2 < v^2 \Leftrightarrow -(z-\bar{z})^2 < 4v^2 \Leftrightarrow 4v^2 + (z-\bar{z})^2 > 0$$

Воспользуемся критерием Сильвестра:

$$\begin{pmatrix} 1 & z-\bar{z} \\ -(z-\bar{z}) & 4v^2 \end{pmatrix} > 0 \Leftrightarrow 1 > 0, 4v^2 + (z-\bar{z})^2 > 0$$

$$L_1 = \begin{pmatrix} 1 & 0 \\ 0 & 4v^2 \end{pmatrix}, M_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

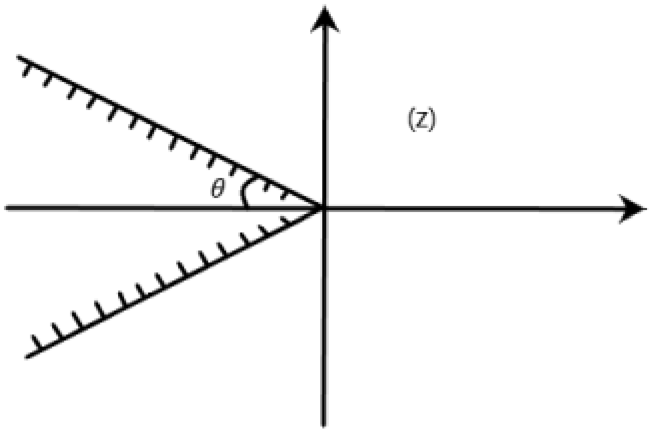
b) $Rez < 0$

$$z + \bar{z} < 0 \Leftrightarrow -z - \bar{z} > 0 \Rightarrow L_2 = 0, M_2 = -1$$

Воспользуемся свойством $LMI$-области об объединении и получим:

$$L = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4v^2 \end{pmatrix}, M = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

4) Внутренность угла: $D = \{z \in \mathbb{C}: |Imz| < -\tan\theta\, Rez, \theta \in [0; 0.5\pi]\}$



$$|Imz| < -\tan\theta\, Rez \Leftrightarrow \left|\frac{z-\bar{z}}{2i}\right|^2 < \tan^2\theta\left(\frac{z+\bar{z}}{2}\right)^2$$

$$-(z-\bar{z})^2 < \tan^2\theta\,(z+\bar{z})^2 \Leftrightarrow \sin^2\theta\,(z+\bar{z})^2 + \cos^2\theta\,(z-\bar{z})^2 > 0$$

Опять-таки применяем критерий Сильвестра:

$$\begin{pmatrix} \sin\theta\,(z+\bar{z}) & \cos\theta\,(z-\bar{z}) \\ -\cos\theta\,(z-\bar{z}) & \sin\theta\,(z+\bar{z}) \end{pmatrix} < 0 \Leftrightarrow$$

$$\Leftrightarrow z+\bar{z} < 0, \qquad \sin^2\theta\,(z+\bar{z})^2 + \cos^2\theta\,(z-\bar{z})^2 > 0$$

$$L = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, M = \begin{pmatrix} \sin\theta & \cos\theta \\ -\cos\theta & \sin\theta \end{pmatrix}$$

Само по себе неравенство $\sin^2\theta\,(z+\bar{z})^2 + \cos^2\theta\,(z-\bar{z})^2$ задаёт область больше, чем показана на рисунке, так как будет ещё симметричная ей область в правой полуплоскости. Это компенсируется неравенством $z+\bar{z} < 0$, которое говорит о том, что мы рассматриваем только левую полуплоскость.

Перед тем как перейти к последней теме этого раздела необходимо вспомнить, а скорее даже узнать, что такое произведение Кронекера, и какими свойствами оно обладает.

*Определение.* Произведение Кронекера пары матриц $A$ и $B$:

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1m}B \\ \dots & \dots & \dots \\ a_{n1}B & \dots & a_{nm}B \end{pmatrix} \in \mathbb{R}^{np\times mq}$$

где

$$A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \in \mathbb{R}^{n\times m}, B = \begin{pmatrix} b_{11} & \dots & b_{1q} \\ \dots & \dots & \dots \\ b_{p1} & \dots & b_{pq} \end{pmatrix} \in \mathbb{R}^{p\times q}$$

Свойства:

1) Отсутствие коммутативности

$$A \otimes B \neq B \otimes A$$

2) Дистрибутивность, ассоциативность, умножение на скаляр

- $(A+B)\otimes C = A\otimes C + B\otimes C,\ C\otimes(A+B) = C\otimes A + C\otimes B$
- $(A\otimes B)\otimes C = A\otimes(B\otimes C)$
- $\alpha(A\otimes B) = (\alpha A)\otimes B = A\otimes(\alpha B)$

3) Смешанное произведение

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$$

Если произведения $AC$ и $BD$ определены.

$$A \otimes B = (I \otimes B)(A \otimes I)$$

$I$ – единичная матрица.

4) Транспонирование

$$(A \otimes B)^T = A^T \otimes B^T$$

5) Обратимость

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$$

Произведение Кронекера обратимо тогда и только тогда, когда обратимы матрицы $A$ и $B$.

6) Сумма и экспонента Кронекера

$$A, I_n \in \mathbb{R}^{n \times n};\ B, I_m \in \mathbb{R}^{m \times m}$$

$I_n, I_m$ – единичные матрицы.

Сумму Кронекера можно определить следующим образом:

$$A \oplus B = A \otimes I_m + I_n \otimes B$$

Экспонента от суммы Кронекера:

$$e^{A \oplus B} = e^A \otimes e^B$$

7) Спектр, след и определитель

$$A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{m \times m}$$

- $\forall k = 1, \dots, n, \forall j = 1, \dots, m$: $\lambda(A \otimes B) = \lambda_k(A)\ \lambda_j(B)$
- $tr(A \otimes B) = tr(A) tr(B)$
- $\det(A \otimes B) = (\det A)^m (\det B)^n$

Да, свойств тут не мало и большая часть из них нам не пригодится, но не лишним всё же будет это знать. А то вдруг зададут дополнительный вопрос на экзамене на эту тему, а вы всё знаете.

Ну что ж, теперь, когда мы знаем, что такое произведение Кронекера, перейдём к ключевой и по совместительству единственной теореме данной главы.

**Теорема.** Критерий $D$-устойчивости матрицы

Матрица $A$ является $D$-устойчивой $\Leftrightarrow \exists X = X^T > 0$:

$$L \otimes X + M \otimes (AX) + M^T \otimes (AX)^T < (>)\ 0,$$

где

$$D = \{z \in \mathbb{C}: F(z) = L + Mz + M^T \bar{z} < (>)\ 0\}$$

Без доказательства.

Это замечательная теорема позволит нам сдвигать собственные числа систем с управлением, в любую нужную область, если это возможно. Именно этим мы и займёмся дальше.

## Задача модального управления относительно $LMI$-области

Итак, мы подошли к кульминации темы линейно-матричных неравенств. Сейчас мы научимся сдвигать собственные числа в заданные $LMI$-области, используя критерий $D$-устойчивости матрицы. Это называется задачей модального управления относительно $LMI$-области.

Мы рассмотрим задачу модального управления для непрерывной и дискретной систем с полной информацией о состоянии. Результаты представим в виде теорем.

**Теорема.**

Система

$$\dot{x} = Ax + Bu, \qquad u = \theta x$$

является асимптотически устойчивой, если

$$\exists X = X^T > 0: XA^T + AX + BZ + Z^T B^T < 0, \qquad Z = \theta X$$

Да, этот результат мы уже получали ранее, только в этот раз докажем это с помощью критерия $D$-устойчивости.

Доказательство:

Непрерывная система является асимптотически устойчивой, если все её собственные числа лежат в левой полуплоскости, то есть:

$$spec\ (A + B\theta) \subseteq int\ D = \{z \in \mathbb{C}: z + \bar{z} < 0\}$$

$D$ – $LMI$-область: $L = 0, M = 1$

Критерий $D$-устойчивости матрицы говорит о том, что для этого необходимо и достаточно, чтобы

$$\exists X = X^T > 0: L \otimes X + M \otimes \big((A + B\theta)X\big) + M^T \otimes \big((A + B\theta)X\big)^T < 0$$

Проведём несколько небольших манипуляций

- Раскрываем внутренние скобки

$$L \otimes X + M \otimes (AX + B\theta X) + M^T \otimes (AX + B\theta X)^T < 0$$

- Делаем замену

$$]Z = \theta X \Rightarrow L \otimes X + M \otimes (AX + BZ) + M^T \otimes (AX + BZ)^T < 0$$

- Подставляем $L$ и $M$

$$XA^T + AX + BZ + Z^T B^T < 0$$

**Теорема.**

Система

$$x_{k+1} = Ax_k + Bu_k, \qquad u_k = \theta x_k$$

является асимптотически устойчивой, если

$$\exists X = X^T > 0: \begin{pmatrix} X & AX + BZ \\ XA^T + Z^T B^T & X \end{pmatrix} > 0, \qquad Z = \theta X$$

Доказательство:

Для асимптотической устойчивости дискретной системы необходимо и достаточно, чтобы все собственные числа лежали внутри единичного круга:

$$spec\ (A + B\theta) \subseteq int\ D = \{z \in \mathbb{C}: |z| < 1\}$$

Как мы знаем, единичный круг – это $LMI$-область:

$$L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Применим критерий $D$-устойчивости для матрицы замкнутой системы:

$$\exists X = X^T > 0: L \otimes X + M \otimes \big((A + B\theta)X\big) + M^T \otimes \big((A + B\theta)X\big)^T > 0$$

Осталось сделать всё то же самое, что и для непрерывной системы.

- Раскрываем внутренние скобки

$$L \otimes X + M \otimes (AX + B\theta X) + M^T \otimes (AX + B\theta X)^T > 0$$

- Делаем замену

$$]Z = \theta X \Rightarrow L \otimes X + M \otimes (AX + BZ) + M^T \otimes (AX + BZ)^T > 0$$

- Подставляем $L$ и $M$

$$\begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix} + \begin{pmatrix} 0 & BZ + AX \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ XA^T + Z^T B^T & 0 \end{pmatrix} > 0$$

Итого получаем то, что было необходимо.

$$\begin{pmatrix} X & AX + BZ \\ XA^T + Z^T B^T & X \end{pmatrix} > 0$$

Ура! Большой блок про линейно-матричные неравенства закончился, а мы плавно подобрались к последней теме – робастная устойчивость.

## Робастная устойчивость. Теорема Харитонова. Теорема Цыпкина-Поляка. Синтез задачи робастной стабилизации при помощи $LMI$

До этого момента мы рассматривали системы, в которых матрицы были постоянными, но часто бывает так, что элементы матриц находятся в некоторых границах, и это приводит нас к новой задаче – задаче робастной стабилизации.

*Задача робастной стабилизации* может быть представлена в виде:

$$\dot{x} = A(\sigma)x + B(\sigma)u, \qquad u = \theta x$$

Необходимо, что система являлась асимптотически устойчивой при любом параметре, то есть

$$\forall\sigma \in D: A(\sigma) + B(\sigma)\theta - \text{асимпт. уст.}$$

Параметр $\sigma$ вызывает так называемую неопределённость.

Эта задача очень сложная и не может быть решена в общем случае, поэтому сейчас мы посмотрим, какие предположения можно сделать, чтобы задача имела разумное решение.

Начнём с того, что рассмотрим несколько видов неопределённостей, но не для матриц, а для многочленов, которыми являются характеристические уравнения этих матриц. Запишем вид этих многочленов:

- Один многочлен

$$P_n(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n$$

- Несколько многочленов $(l + 1)$

$$\forall k = 0, \dots, l: P_{n,k}(z) = a_{0,k} + a_{1,k} z + a_{2,k} z^2 + \cdots + a_{n,k} z^n$$

*Определение.* Многочлен $P_n(z)$ называется интервальным, если

$$\forall k = 0, \dots, n: \underline{a}_k \le a_k \le \overline{a}_k$$

Неопределённость, которая его задаёт, называют *интервальной*.

*Определение.* Многочлен вида

$$P_n(z, q) = P_{n,0}(z) + P_{n,1}(z)q_1 + \cdots + P_{n,l}(z)q_l, \qquad q = (q_1 \quad \cdots \quad q_l)^T \in D$$

называют аффинным, а неопределённость, которая его задаёт, называют *аффинной*, как не трудно было догадаться.

На этом этапе уже есть интересные результаты, одним из которых является теорема Харитонова для интервального многочлена.

**Теорема Харитонова.**

Интервальный многочлен

$$P_n(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n, \qquad \underline{a}_k \le a_k \le \overline{a}_k$$

робастно устойчив, если устойчивы многочлены Харитонова:

$$\underline{a}_0 + \overline{a}_1 z + \overline{a}_2 z^2 + \underline{a}_3 z^3 + \underline{a}_4 z^4 + \cdots = 0$$

$$\underline{a}_0 + \underline{a}_1 z + \overline{a}_2 z^2 + \overline{a}_3 z^3 + \underline{a}_4 z^4 + \cdots = 0$$

$$\overline{a}_0 + \underline{a}_1 z + \underline{a}_2 z^2 + \overline{a}_3 z^3 + \overline{a}_4 z^4 + \cdots = 0$$

$$\overline{a}_0 + \overline{a}_1 z + \underline{a}_2 z^2 + \underline{a}_3 z^3 + \overline{a}_4 z^4 + \cdots = 0$$

Без доказательства.

Да, эта теорема позволяет свести проверку бесконечного множества многочленов на устойчивость до проверки всего четырёх. Прекрасный результат.

Проверять многочлены на устойчивость можно по-разному. Рекомендую следующий вариант.

- До четвёртой степени использовать критерий Рауса-Гурвица
- После четвёртой метод лямбда-перехода

Теорема Харитонова хороша, но она позволяет лишь ответить на вопрос устойчив ли интервальный многочлен в заданных границах, но не позволяет узнать границы устойчивости. Как раз этот изъян отсутствует в следующей теореме.

Допустим нам известен всё тот же интервальный многочлен

$$P_n(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n$$

но его границы имеют несколько иной вид:

$$\forall k = 0, \dots n \colon \left| a_k - a_k^0 \right| < \gamma \alpha_k$$

тогда это уже не просто многочлен, а целое семейство многочленов, в котором $\gamma$ выступает параметром.

Теперь всё готово, чтобы записать теорему.

**Теорема Цыпкина-Поляка.**

Интервальное семейство многочленов робастно устойчиво тогда и только тогда, когда

- $a_0^0 > \gamma\alpha_0$
- $a_n^0 > \gamma\alpha_n$
- Годограф Цыпкина-Поляка при изменении $\omega$ от нуля до бесконечности последовательно проходит $n$ квадрантов против часовой стрелки и не пересекает квадрат с вершинами в точках $(\pm\gamma, \pm\gamma)$

Радиус робастной устойчивости:

$$\gamma_{max} = \min\left\{\gamma^*, \frac{a_0^0}{\alpha_0}, \frac{a_n^0}{\alpha_n}\right\}$$

$\gamma^*$ - максимальное значение, при котором годограф не пересекается с квадратом $(\pm\gamma, \pm\gamma)$

Это частотный критерий, поэтому интервальный многочлен после подстановки $z = i\omega$ примет вид:

$$P_n(i\omega) = u_0(\omega) + i\omega v_0(\omega)$$

Годограф Цыпкина-Поляка имеет вид:

$$\left\{\frac{u_0(\omega)}{R(\omega)}, \frac{v_0(\omega)}{T(\omega)}\right\}$$

где

$$u_0(\omega) = a_0^0 - a_2^0\omega^2 + a_4^0\omega^4 - \cdots, \qquad v_0(\omega) = a_1^0 - a_3^0\omega^2 + a_5^0\omega^4 - \cdots$$

$$R(\omega) = \alpha_0 + \alpha_2\omega^2 + \alpha_4\omega^4 + \cdots, \qquad T(\omega) = \alpha_1 + \alpha_3\omega^2 + \alpha_5\omega^4 + \cdots$$

Без доказательства.

Отлично, теперь мы можем найти границы робастной устойчивости! Но это всё только для интервальной неопределённости. К сожалению, на этом задачи робастной устойчивости не заканчиваются, так что мы движемся дальше.

Предположим, что матрицы $A(\sigma)$ и $B(\sigma)$ являются симплексами, то есть

$$A(\sigma) = \sum_{i=1}^{m} A_k\sigma_k, B(\sigma) = \sum_{i=1}^{m} B_k\sigma_k: \ \sigma_k \in [0; 1], \sum_{i=1}^{m} \sigma_k = 1$$

тогда система

$$\dot{x} = A(\sigma)x + B(\sigma)u, \qquad u = \theta x$$

может быть стабилизирована.

**Теорема.** Достаточное условие робастной стабилизируемости.

Система робастно стабилизируема, если

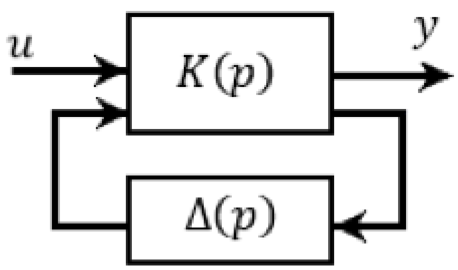
$$\exists X = X^T > 0, Z = \theta X: \forall k = 1, \dots, m \to XA_k^T + A_k X + B_k Z + Z^T B_k^T < 0$$

Без доказательства.

Это лишь достаточное условие, так что оно не гарантирует нам, что система не стабилизируема, если у этого линейно-матричного неравенства нет решения.

Бонус для тех, кто дочитал, - ещё одно определение неопределённости!

*Определение.* Структурной неопределённостью называется $\Delta(p)$:




Ну, на этом всё. Вы прошли этот курс! Строчка за строчкой, страничка за страничкой. Вы справились, поздравляю!

## Приложение

**Утверждение.** Тождество Шермана-Моррисона-Вудбери

$$(A + XRY)^{-1} = A^{-1} - A^{-1}X(R^{-1} + YA^{-1}X)^{-1}YA^{-1}$$

Доказательство:

$]T = (R^{-1} + YA^{-1}X)$

$$\begin{aligned}(A + XRY)(A + XRY)^{-1} &= (A + XRY)(A^{-1} - A^{-1}XT^{-1}YA^{-1}) \\ &= I + XRYA^{-1} - XT^{-1}YA^{-1} - XRYA^{-1}XT^{-1}YA^{-1}\end{aligned}$$

- Заметим, что

$$X(I + RYA^{-1}X)T^{-1}YA^{-1} = XRR^{-1}(I + RYA^{-1}X)T^{-1}YA^{-1} = XRYA^{-1}$$

- Применяем это и получаем ответ

$$I + XRYA^{-1} - X(I + RYA^{-1}X)T^{-1}YA^{-1} = I + XRYA^{-1} - XRYA^{-1} = I$$

**Утверждение.** Математическое ожидание квадратичной формы.

$$M\{x^T Sx\} = \bar{x}^T S\bar{x} + tr\{SX\}, \qquad M\{x\} = \bar{x}, M\{(x - \bar{x})(x - \bar{x})^T\} = X$$

Доказательство:

$$M\{x^T Sx\} = M\big\{tr\{x^T Sx\}\big\}$$

- Воспользуемся свойством следа и продвинем мат. ожидание внутрь

$$M\big\{tr\{Sxx^T\}\big\} = tr\big\{SM\{xx^T\}\big\}$$

- Вспомним связь ковариационной и корреляционной матриц

$$M\{(x - \bar{x})(x - \bar{x})^T\} = M\{xx^T\} - M\{x\}M^T\{x\} \Rightarrow M\{xx^T\} = X + \bar{x}\bar{x}^T$$

- Ещё раз воспользуемся свойством следа

$$tr\big\{SM\{xx^T\}\big\} = tr\{SX + S\bar{x}\bar{x}^T\} = tr\{SX\} + \bar{x}^T S\bar{x}$$

## Литература и другие ресурсы

1) Лекции по курсу

Лекций нигде нет, но оставлю ссылку на e-learning, там есть полезная информация

https://e-learning.unn.ru/course/view.php?id=1515

2) Критерий Сильвестра

https://ru.wikipedia.org/wiki/Критерий_Сильвестра

3) Линейно-матричные неравенства

https://ru.wikipedia.org/wiki/Линейные_матричные_неравенства

4) Векторное и матричное дифференцирование

https://ml-handbook.ru/chapters/matrix_diff/intro

http://nabatchikov.com/blog/view/matrix_der

https://translated.turbopages.org/proxy_u/en-ru.ru.a2bca5b9-6274c4f7-4099e7b0-74722d776562/https/en.wikipedia.org/wiki/Matrix_differentiation

5) След матрицы

https://ru.wikipedia.org/wiki/След_матрицы

6) Условные величины

https://brilliant.org/wiki/conditional-probability-distribution/

https://ru.wikipedia.org/wiki/Условное_математическое_ожидание

7) Спонсор хорошего настроения: аниме “Волейбол!”, “Необъятный океан”, “Ванпанчмэн”, “Почувствуй ветер”