ITMO UNIVERSITY

Saint Petersburg, Russia

Математические основы теории систем

Линейные и квадратичные формы

2022

В этой лекции обсудим понятия:

1) Скалярной функции от вектора,
2) Линейных и квадратичных форм,
3) Положительной и отрицательной определенности квадратичной формы,
4) Правила дифференцирования функций от векторов и матриц по скалярным, векторным и матричным переменным

Определение 1 (скалярная функция от вектора). Поставим в соответствие каждому вектору линейного действительного пространства R^n вполне определенное число из R. В этом случае говорят, что в линейном пространстве R^n определена *скалярная* функция F от вектора x :

$$F(x): R^n \to R.$$

Определение 2 (действительная линейная форма) Функция $F_1(x)$ областью определения которой является линейное пространство R^n а областью значений – совокупность действительных чисел R называется действительной *линейной формой* (линейным функционалом), если выполняется соотношение

$$F_1(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 F_1(x_1) + \alpha_2 F_1(x_2) \quad (1)$$

для любых векторов x_1 и x_2 и любых действиельных чисел α_1 и α_2.

Линейная форма – это скалярная функция векторного аргумента

Рассмотрим:
$\{e_1, e_2, \dots, e_n\}$ – естественный базис в пространстве R^n, $x = [x_1, x_2, \dots, x_n]^T$ – координаты вектора x в этом базисе.
Тогда линейная форма может быть представлена как

$$F_1(x) = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n \,, \qquad (2)$$

где $\alpha_k = F_1(e_k), k = 1, \dots, n$.
И наоборот, при любых действительных числах α_1, α_2, …, α_n выражение (2) определяет некоторую линейную форму в R^n.
Определение 3 (ядро линейной формы). Множество всех векторов $x \in R^n$, для которых $F_1(x) = 0$, называется *ядром* линейной формы (функционала) и обозначается $N(F_1)$:

$$N(F_1) = \{x \in R^n \colon F_1(x) = 0\} \qquad (3)$$

Линейную форму (2) можно записать как скалярное произведение векторов x и α :

$$F_1(x) = (x, \alpha) = (\alpha, x)$$

Перейдем к рассмотрению квадратичных форм.

Определение 4 (квадратичная форма). Квадратичной формой от n действительных переменных $x_1, x_2, \dots, x_n$ называется функция вида

$$F_2(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j, \quad (4)$$

где a_{ij} – действительные числа.

Матрица $A = [a_{ij}]$ – матрица квадратичной формы. Это симметричная матрица.

Квадратичная форма – это скалярная функция векторного аргумента.

Пример.

Рассмотрим квадратичную форму $F_2(x) = 2x_1^2 - 3x_1x_2 + 4x_1x_3 + 5x_2^2 - 8x_2x_3 + x_3^2$.

Для того, чтобы составить матрицу A, нам нужно представить слагаемые $-3x_1x_2, 4x_1x_3$ и $-8x_2x_3$ в виде сумм двух равных слагаемых:

$$-3x_1x_2 = -1.5x_1x_2 - 1.5x_2x_1,$$

$$4x_1x_3 = 2x_1x_3 + 2x_3x_1,$$

$$-8x_2x_3 = -4x_2x_3 - 4x_3x_2.$$

Теперь можно записать матрицу $A = [a_{ij}]$

$$A = \begin{bmatrix} 2 & -1.5 & 2 \\ -1.5 & 5 & -4 \\ 2 & -4 & 1 \end{bmatrix}.$$

Квадратичная форма может быть записана как

$$F_2(x) = x^T Ax. \quad (5)$$

Ранг квадраичной формы – это ранг ее матрицы A.

Преобразование матрицы квадратичной формы при линейном преобразовании координат

Рассмотрим линейное преобразование координат y в x:

$$x_1 = t_{11}y_1 + t_{12}y_2 + \cdots + t_{1n}y_n$$

$$x_2 = t_{21}y_1 + t_{22}y_2 + \cdots + t_{2n}y_n$$

$$\vdots$$

$$x_n = t_{n1}y_1 + t_{n2}y_2 + \cdots + t_{nn}y_n$$

Эта замена координат может быть записана как

$$x = Ty.$$

Для новых координат y квадратичная форма выглядит следующим образом:

$$F_2(y) = (Ty)^T A(Ty) = y^T T^T ATy = y^T By.$$

Матрицы A и B связаны соотношением $B = T^T AT$.

Канонический вид квадратичной формы

Любую квадратичную форму $F_2(x)$ можно привести к каноническому виду невырожденным преобразованием $x = Ty$:

$$F_2(y) = \sum_{i=1}^{n} \lambda_i (y_i)^2 . \qquad (6)$$

Матрица квадратичной формы $F_2(y)$ – это диагональная матрицв с λ_i на главной диагонали.

Квадратичная форма может быть приведена к каноническому виду с помощью ортогональных преобразований. Преобразование координат $x = Ty$ называется ортогональным, если его матрица T ортогональная.

Для любой квадратичной формы $F_2(x) = x^T Ax$ существует ортогональное преобразование $x = Ty$ которое приводит ее к каноническому виду

$$F_2(y) = \sum_{i=1}^{n} \lambda_i (y_i)^2 .$$

Здесь λ_i $(i = 1, \dots n)$– собственные числа матрицы A, столбцы матрицы T – попарно ортогональные собственные векторы A (норма каждого собственного вектора равна 1).

Пример. Рассмотрим матрицу квадратичной формы A

$$A = \begin{bmatrix} 3/2 & 0 & -1/2 \\ 0 & 1 & 0 \\ -1/2 & 0 & 3/2 \end{bmatrix}.$$

Используем ортогональное преобразование для того, чтобы привести ее к каноническому виду $\Lambda = T^T AT$.

Решаем пример по шагам:

1) Вычислим собственные значения матрицы A:

$$\det(\lambda I - A) = 0,\ \lambda_1 = \lambda_2 = 1,\ \lambda_3 = 2$$

Как видно, кратность $\lambda = 1 \quad m = 2$.

2) Вычислим собственные векторы $\xi_i = \begin{bmatrix} \xi_{i1} \\ \xi_{i2} \\ \xi_{i3} \end{bmatrix}, i = 1, 2, 3$ матрицы A: $(A - I\lambda_i)\xi_i = 0$

Для $\lambda_1 = \lambda_2 = 1$ имеем следующую систему уравнений:

$$\left(\begin{bmatrix} 3/2 & 0 & -1/2 \\ 0 & 1 & 0 \\ -1/2 & 0 & 3/2 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} \xi_{i1} \\ \xi_{i2} \\ \xi_{i3} \end{bmatrix} = 0, i = 1,2$$

ITMOre than a UNIVERSITY

$$\begin{bmatrix} 1/2 & 0 & -1/2 \\ 0 & 0 & 0 \\ -1/2 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} \xi_{i1} \\ \xi_{i2} \\ \xi_{i3} \end{bmatrix} = 0$$

$$\tfrac{1}{2}\xi_{i1} - \tfrac{1}{2}\xi_{i3} = 0$$

$$0 \cdot \xi_{i1} + 0 \cdot \xi_{i2} + 0 \cdot \xi_{i3} = 0$$

$$-\frac{1}{2}\xi_{i1} + \frac{1}{2}\xi_{i3} = 0$$

Выберем векоры ξ_1 и ξ_2 так, чтобы они были линейно независимы и ортонормальны ($\xi_{i1} = \xi_{i3}$, ξ_{i2} принимает любые значения, $i = 1, 2$):

$$\xi_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{bmatrix}, \qquad \xi_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Для $\lambda_3 = 2$ имеем:

$$\left(\begin{bmatrix} 3/2 & 0 & -1/2 \\ 0 & 1 & 0 \\ -1/2 & 0 & 3/2 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}\right)\begin{bmatrix} \xi_{i1} \\ \xi_{i2} \\ \xi_{i3} \end{bmatrix} = 0, i = 3$$

$$-\frac{1}{2}\xi_{i1} - \frac{1}{2}\xi_{i3} = 0$$

$$-\xi_{i2} = 0$$

$$-\frac{1}{2}\xi_{i1} + \frac{1}{2}\xi_{i3} = 0$$

Поэтому выбираем $\xi_3 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ -\frac{\sqrt{2}}{2} \end{bmatrix}$.

Тогда матрица ортогонального преобразования T равна

$$T = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \end{bmatrix}.$$

И каноническая форма $\Lambda = T^T AT$

$$\Lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Положительная и отрицательная определенность квадратичных форм

Запишем квадратичную форму как скалярное произведение с весом A:

$$F_2(x) = (Ax, x) = x^T Ax = (\|x\|)_A^2$$

Определение 5 **(Положительная и отрицательная определенность).** Квадратичная форма $F_2(x) = (Ax, x)$ называется положительно (отрицательно) определенной, если $(Ax, x) > 0$ $((Ax, x) < 0)$ для $x \neq 0$ и $(Ax, x) = 0$ для $x = 0$.
Квадратичная форма $F_2(x) = (Ax, x)$ называется неотрицательно (неположительно) определенной, если она принимает только неотрицательные (неположительные) значения и при этом обращается в нуль не только для $x = 0$.
Если квадратичная форма $F_2(x)$ принимает как положительные, так и отрицательные значения, то она называется неопределенной (или знакопеременной).

Как определить, является ли квадратичная форма положительно определенной?

Теорема 1. Если все собственные значения матрицы квадратичной формы являются положительными, то тогда квадратичная форма – положительно определенная. Если все собственные значения, то тогда квадратичная форма – отрицательно определенная.

Пример. Вернемся к матрице $A = \begin{bmatrix} 3/2 & 0 & -1/2 \\ 0 & 1 & 0 \\ -1/2 & 0 & 3/2 \end{bmatrix}$ с собственными числами $\lambda_1 = \lambda_2 = 1,\ \lambda_3 = 2.$ Все они положительные $\Longrightarrow$ квадратичная форма – положительно определенная.

Также для того, чтобы проверить положительную определенность, можно применить критерий Сильвестра.

Теорема 2 (Критерий Сильвестра).

Рассмотрим угловые миноры матрицы A:

$$\delta_1 = a_{11}, \quad \delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \dots, \quad \delta_k = \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} \end{vmatrix}, \quad \dots$$

$$\dots, \quad \delta_n = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$$

Для того, чтобы квадратичная форма была положительно определенной, необходимо и достаточно, чтобы все угловые миноры ее матрицы были положительны.

Для того, чтобы квадратичная форма была отрицательно определенной, необходимо и достаточно, чтобы знаки угловых миноров ее матрицы чередовались следующим образом:

$$\delta_1 < 0, \delta_2 > 0, \delta_3 < 0, \dots$$

Пример.

a) Рассмотрим квадратичную форму $F_2(x) = -x_2^2 + 4x_1x_2$ с матрицей $A = \begin{bmatrix} 0 & 2 \\ 2 & -1 \end{bmatrix}$

Угловые миноры $\delta_1 = 0$, $\delta_2 = \begin{vmatrix} 0 & 2 \\ 2 & -1 \end{vmatrix} = -4 < 0 \Rightarrow$ *форма не знакоопределенная*

b) Рассмотрим квадратичную форму $F_2(x) = 3x_1^2 - 2x_2^2 + 4x_1x_2 - x_3^2 + 2x_2x_3$ с матрицей $A = \begin{bmatrix} 3 & 2 & 0 \\ 2 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$

$\delta_1 = 3 \Rightarrow$ *форма точно не отрицательно определенная*

$\delta_2 = \begin{vmatrix} 3 & 2 \\ 2 & -2 \end{vmatrix} = -6 - 4 < 0 \Rightarrow$ *форма не положительно определенная*

Заключение: форма не знакоопределенная

Задание

Определите, является ли эта квадратичная форма знакоопределенной

$$F_2(x) = x_2^2 - 4x_1x_2 + 2x_2x_3$$

Правила дифференцирования функций от векторов и матриц по скалярным, векторным и матричным переменным

Определение 6 (производная от матрицы по скалярной переменной). Пусть A – это матрица, элементами которой являются функции $A_{ij} = A_{ij}(q)$ скалярной переменной q . Тогда произволная матрицы $A(q)$ это матрица $A_q = \frac{\partial A(q)}{\partial q}$, составленная из производных ее элементов $A_{ij_q} = \frac{\partial A_{ij}(q)}{\partial q}$ по переменной q, which что может быть записано в форме

$$A_q = row\{col(A_{ij}; i = 1, \dots m); j = 1, \dots n\}$$

Пример. **Производная от матрицы по скалярной переменной**

$$A(q) = \begin{bmatrix} 0 & 1 \\ -10(1+q^2) & -7(1+q) \end{bmatrix}$$

$$A_q = \frac{\partial A(q)}{\partial q} = \begin{bmatrix} 0 & 0 \\ -20q & -7 \end{bmatrix}$$

Определение 7 (Производная скалярной функции по векторному аргументу). Пусть $J = J(x)$ – скалярная функция векторного аргумента $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. Тогда, обозначив символом ∇ оператор градиента, для *производной* этой функции $\frac{\partial J}{\partial x}$ *по векторному аргументу* x и градиента можно записать следующие представления:

$$\nabla_x J = \left(\frac{\partial J}{\partial x}\right)^T = \left[\frac{\partial J}{\partial x_1}, \frac{\partial J}{\partial x_2}, \dots, \frac{\partial J}{\partial x_n}\right]^T$$ - вектор–столбец

$$(\nabla_x J)^T = \frac{\partial J}{\partial x} = \left[\frac{\partial J}{\partial x_1}, \frac{\partial J}{\partial x_2}, \dots, \frac{\partial J}{\partial x_n}\right]$$ - вектор–строка

$$\nabla_{xx} J = \frac{\partial}{\partial x} \cdot \frac{\partial J}{\partial x} = \begin{bmatrix} \frac{\partial^2 J}{\partial x_1^2} & \cdots & \frac{\partial^2 J}{\partial x_n \partial x_1} \\ \cdots & \cdots & \cdots \\ \frac{\partial^2 J}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 J}{\partial x_n^2} \end{bmatrix}$$ $(n \times n)$ - матрица

Определение 8 (производная векторной функции по векторному аргументу).

Пусть $y = \begin{bmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_m(x) \end{bmatrix}$ – вектор скалярных функций векторного аргумента $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ (векторная функция от вектора). Тогда

$$\frac{dy(x)}{dx} = \begin{bmatrix} \frac{\partial y_1(x)}{\partial x_1} & \cdots & \frac{\partial y_1(x)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m(x)}{\partial x_1} & \cdots & \frac{\partial y_m(x)}{\partial x_n} \end{bmatrix} - m \times n \text{ –матрица.}$$

Пример (производная векторной функции по векторному аргументу).

$$y(x) = \begin{bmatrix} y_1(x) \\ y_2(x) \end{bmatrix} = \begin{bmatrix} x_1^2 x_2 \\ x_2 \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\frac{dy}{dx} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 x_2 & x_1^2 \\ 0 & x_2 \end{bmatrix}$$

Thank you for your attention!

www.ifmo.ru

ITsMOre than a UNIVERSITY