Chapter 1

Многомерная оптимизация функции многих переменных

Определение: Произвольное n-мерное пространство будем называть евклидовым, если мы на нем введем скалярное произведение.

$$(a,b) = (b,a)$$
$$(\lambda a, b) = \lambda(a,b)$$
$$(a,a) \geq 0; (a,a) = 0 \Leftrightarrow a = 0$$

Определение: Нормированным пространством будем называть евклидово пространство с введеной на нем нормой вида

$$||x|| = \sqrt{(x,x)}$$

Свойства

$$||x|| \geq 0; ||x|| = 0 \Leftrightarrow x = 0$$
$$||\lambda x|| = |\lambda| \cdot ||x||$$
$$||x+y|| \leq ||x|| + ||y||$$

Теорема: Всякое скалярное умножение в евклидовом пространстве определяется по формуле

$$||x|| = \sqrt{(x,x)} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Если менять векторы на квадратные матрицы, то добавится еще аксиома:

$$||AB|| \leq ||A|| \cdot ||B||$$

Норма матрицы - неотрицательное число.

Из последней аксиомы, если взять вместо одной матрицы вектор, можно получить

$$||A|| = \sup_{||x||\neq 0} \frac{||Ax||}{||x||}$$

- из этого условия будет следовать условие согласованности нормы матрицы и нормы вектора, т.е.

$$||Ax|| \leq ||A|| \cdot ||x||$$

Определение: функция f будет называться квадратичной формой, если

$$f(\overline{x}) = f(x_1, x_2, ..., x_n) = \sum_{i=1;j=1}^{n} a_{i,j}x_i x_j$$

Можно привести к виду

$$f(\overline{x}) = ((\overline{x})^T, A\overline{x})$$

или можно заменой

$$\sum_{i=1;j=1}^{n} a_{i,j}x_i x_j = \sum_{i=1}^{n} q_i^2 + 2\sum_{i,j=1}^{n} b_{i,j}q_i q_j$$

Определение: матрицы A будем называть неотрицательно определенными, если $\forall \overline{x} \in E^n((\overline{x})^T, A\overline{x}) \geq 0$, положительно - если > 0, и отрицательно - если < 0

Диагональный (нормальный) вид матрицы - с собственными значениями на диагонали, т.е. для положительно определенной

$$((\overline{x})^T, A\overline{x}) = \sum_{i=1}^{n} \lambda_i x_i^2$$

Канонический - с единицами на диагонали, т.е. заменой с нормального вида для положительного

$$y_i = \sqrt{\lambda_i} x_i$$

Для отрицательного общий вид

$$(-1)^n \sum_{i=1}^{n} y_i^2$$

Теорема: симметричная матрица A является положительно определенной когда все определители главных миноров матрицы A неотрицательны. (Критерий Сильвестра)

Матрица будет отрицательно определенной, если знаки определителей главных миноров чередуются.

Определение: Линией уровня будем называть множество точек $f(\overline{x}) =$ Const $= C$ например $f(x_1, x_2) = x_1^2 + 4x_2^2 = 4$ - эллипс.

1.1 Градиент функции и матрица Гессе

Пусть задана функция $f : E^n \to \mathbb{R}$ - наша задача найти миниммы функции в E^n

Определение: градиентом функции в точке x^* будет

$$\nabla f(x^*) = \left(\frac{\partial f(x^*)}{\partial x_1}, ..., \frac{\partial f(x^*)}{\partial x_n}\right)$$

Точка $x^* \in E^n$ в которой $\nabla f(x^*) = 0$ будет являться стационарной точкой.

Матрицей Гессе будем называть матрицу вторых производных функции

$$f''(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & ... & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ ... & ... & ... & ... \\ ... & ... & ... & ... \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & ... & ... & ... \end{pmatrix} = H$$

(она симметричная)

Теорема: (Необходимое условие экстремума функции) Пусть дана $f : E^n \to \mathbb{R}$ (действует из евклидового пространства во множество вещественных чисел) - она дифференцируемая в точке $x^* \in E^n$ - тогда если x^* - это точка безусловного локального экстремума, то градиент функции в этой точке равен нулю.

Теорема: (Необходимое условие экстремума второго порядка) Пусть та же функция f уже дважды дифференцируема в x^* - если x^* - точка безусловного локального минимума функции $f(x)$ то матрица Гессе в точке x^* будет неотрицательно определена. Если это точка локального максимума, то матрица Гессе будет неположительно определена.