ДЗ 5 по Методам Оптимизации. Векторное дифференцирование

Соколов Игорь, группа 573

8 октября 2017 г.

1

Решение:

Ответ:

2

Найти $\nabla f(x), f''(x)$, если $f(x) = \dfrac{-1}{1 + x^T x}$

Решение:

Перепишем $f(x)$ в скаляром виде:

$$f(x) = \frac{-1}{1 + \sum\limits_{i=1}^{n} x_i^2}$$

$$\frac{\partial f(x)}{\partial x_k} = \frac{\partial}{\partial x_k}\left(\frac{-1}{1 + \sum\limits_{i=1}^{n} x_i^2}\right) = \frac{2x_k}{\left(1 + \sum\limits_{i=1}^{n} x_i^2\right)^2}$$

$$\Rightarrow \nabla f(x) = \frac{2x}{(1 + x^T x)^2}$$

Пусть $g_k = \dfrac{\partial f(x)}{\partial x_k} \to H_{k,p} = \dfrac{\partial g_k}{\partial x_p}$

Тогда

$$H_{k,p} = \frac{\partial^2 f(x)}{\partial x_p \partial x_k} = \frac{\partial}{\partial x_p}\left(\frac{2x_k}{\left(1+\sum\limits_{i=1}^{n} x_i^2\right)^2}\right) = \frac{2\delta_{k,p}\left(1+\sum\limits_{i=1}^{n} x_i^2\right)^2 - 8x_k x_p\left(1+\sum\limits_{i=1}^{n} x_i^2\right)}{\left(1+\sum\limits_{i=1}^{n} x_i^2\right)^4} =$$

$$= \frac{2}{\left(1+\sum\limits_{i=1}^{n} x_i^2\right)^2}\left(\delta_{k,p} - \frac{4x_p x_k}{1+\sum\limits_{i=1}^{n} x_i^2}\right) \quad (1)$$

$$\Rightarrow f''(x) = \frac{2}{(1+x^T x)^2}\left(\mathbb{E} - \frac{xx^T}{1+x^T x}\right)$$

Ответ:
$\nabla f(x) = \dfrac{2x}{(1+x^T x)^2}$
$f''(x) = \dfrac{2}{(1+x^T x)^2}\left(\mathbb{E} - \dfrac{xx^T}{1+x^T x}\right)$

3

Найти $f'(X)$, если $f(X) = \det X$

Решение:

Пусть $g(X) = \log\det X$
Найдем $g(X+\Delta X) = \log\det X$, где ΔX - некоторая малая матрица по норме Фробениуса.
Тогда $f(X+\Delta X) = e^{g(X+\Delta X)} = \det(X+\Delta X)$

Вспомним, что производную имеет смысл линейной аппроксимация функции в окрестности точки.
Заметим, что:

$$\begin{aligned}\log\det[X+\Delta X] &= \log\det\left[X^{1/2}\left(I + X^{-1/2}\Delta X X^{-1/2}\right)X^{1/2}\right] = \\ &= \log\det\left[X^{1/2}\right]\det\left[I + X^{-1/2}\Delta X X^{-1/2}\right]\det\left[X^{1/2}\right] = \\ &= \log\det[X]\det\left[I + X^{-1/2}\Delta X X^{-1/2}\right] = \\ &= \log\det[X] + \log\det\left[I + X^{-1/2}\Delta X X^{-1/2}\right] \quad (1)\end{aligned}$$

Известно, что определитель матрицы равен произведению её собственных значений

$$\log\det[X+\Delta X] = \log\det X + \sum_{i=1}^{n}\log(1+\lambda_i)$$

Где λ_i - собственные числа матрицы $X^{-1/2}\Delta X X^{-1/2}$. Далее используем факт "малости"матрицы ΔX (в смысле нормы этой матрицы), следовательно для приближения первого порядка справедливо: $\log(1+\lambda_i) \approx \lambda_i$ т.к. λ_i так же должны быть малыми.

$$\log\det\left[X+\Delta X\right] \approx \log\det X + \sum_{i=1}^{n} \lambda_i \log\det\left[X+\Delta X\right] \approx \log\det X + \mathbf{tr}\left[X^{-1/2}\Delta X X^{-1/2}\right] =$$
$$= \log\det X + \mathbf{tr}\left[X^{-1/2}X^{-1/2}\Delta X\right] = \log\det X + \mathbf{tr}\left[X^{-1}\Delta X\right] \quad (2)$$

Заметим, что след произведения матриц есть обобщения скалярного произведения на пространство матриц.

След матрицы обладает следующий свойством:

$$\mathbf{tr}(A^T B) = \mathbf{tr}(AB^T) = \mathbf{tr}(B^T A) = \mathbf{tr}(BA^T)$$

.

Стало быть, справедливо:

$$g(X+\Delta X) \approx g(X) + \langle g'(X), \Delta X\rangle \quad (3)$$

Из (2) $\Rightarrow$

$$e^{\log\det[X+\Delta X]} = e^{\log\det X + \mathbf{tr}\left[X^{-1}\Delta X\right]}$$
$$\det\left[X+\Delta X\right] = e^{\log\det X} e^{\mathbf{tr}\left[X^{-1}\Delta X\right]}$$

В силу малости матрицы ΔX справедливо $e^{\mathbf{tr}\left[X^{-1}\Delta X\right]} = 1 + \mathbf{tr}\left[X^{-1}\Delta X\right]$

$$\Rightarrow \det\left[X+\Delta X\right] = \det X\left(1 + \mathbf{tr}\left[X^{-1}\Delta X\right]\right)$$
$$\det\left[X+\Delta X\right] = \det X + \det X\,\mathbf{tr}\left[X^{-1}\Delta X\right] \quad (4)$$

Заметим, что $detX$ - скаляр и в силу свойства линейности следа матрицы

$$\mathbf{tr}\left(\alpha A\right) = \alpha\mathbf{tr}\left(A\right)$$

Получаем

$$\det\left[X+\Delta X\right] = \det X + \mathbf{tr}\left[(\det X)X^{-1}\Delta X\right]$$

Пользуясь аналогичными примеру (3) рассуждениями, получаем:

$$\det\left[X+\Delta X\right] = \det X + \mathbf{tr}\langle(\det X)X^{-1}, \Delta X\rangle$$

Значит, $f'(X) = det(X)X^{-1}$

Ответ: $f'(X) = det(X)X^{-1}$

4

Найти $f''(X)$, если $f(X) = \log \det X$

Решение:

Ответ:

5

Найти градиент и гессиан функции $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \log \sum\limits_{i=1}^{m} \exp(a_i^T x + b_i), \quad a_1, \ldots, a_m \in \mathbb{R}^n; \quad b_1, \ldots, b_m \in \mathbb{R}$

Решение:

Ответ:

6

Решение: