Ахмаджонов Мумтозбек Б05-927б

Домашнее задание №2

Задача 1.

Найти ∇f, где $f(\overline{x}) = ||A\overline{x}|| - ||\overline{x}^T A||$.

Решение:

Пусть

$$A = \begin{bmatrix} \overline{a_1} \\ \overline{a_2} \\ \vdots \\ \overline{a_n} \end{bmatrix} = \begin{bmatrix} \overline{b_1} & \overline{b_2} & \dots & \overline{b_n} \end{bmatrix},\ A_{ij} = a_{ij} = b_{ji}$$

Тогда

$$||A\overline{x}|| = \sqrt{\sum\nolimits_{i=1}^{n} \left(\overline{a_i}, \overline{x}\right)^2} = \sqrt{\sum\nolimits_{i=1}^{n} \left(\sum\nolimits_{j=1}^{n} a_{ij}x_j\right)^2}$$

$$\frac{\partial ||A\overline{x}||}{\partial x_k} = \frac{1}{2} \frac{\sum\nolimits_{i=1}^{n} 2\left(\sum\nolimits_{j=1}^{n} a_{ij}x_j\right)a_{ik}}{\sqrt{\sum\nolimits_{i=1}^{n} \left(\sum\nolimits_{j=1}^{n} a_{ij}x_j\right)^2}} \Longrightarrow \nabla||A\overline{x}|| = \frac{A^T A\overline{x}}{||A\overline{x}||}$$

Теперь вычислим $\nabla||\overline{x}^T A||$:

$$||\overline{x}^T A|| = \sqrt{\sum\nolimits_{i=1}^{n} \left(\overline{x}, \overline{b_i}\right)^2} = \sqrt{\sum\nolimits_{i=1}^{n} \left(\sum\nolimits_{j=1}^{n} x_j b_{ij}\right)^2}$$

$$\frac{\partial ||\overline{x}^T A||}{\partial x_k} = \frac{1}{2} \frac{\sum\nolimits_{i=1}^{n} 2\left(\sum\nolimits_{j=1}^{n} x_j b_{ij}\right)b_{ik}}{\sqrt{\sum\nolimits_{i=1}^{n} \left(\sum\nolimits_{j=1}^{n} x_j b_{ij}\right)^2}} = \frac{\sum\nolimits_{i=1}^{n} \left(\overline{x}, \overline{b_i}\right)b_{ik}}{\sqrt{\sum\nolimits_{i=1}^{n} \left(\sum\nolimits_{j=1}^{n} x_j b_{ij}\right)^2}} \Longrightarrow$$

$$\Longrightarrow \nabla||\overline{x}^T A|| = \frac{\overline{x}^T A A^T}{||\overline{x}^T A||}.$$

Ответ: $\nabla f(\overline{x}) = \dfrac{A^T A\overline{x}}{||A\overline{x}||} - \dfrac{\overline{x}^T A A^T}{||\overline{x}^T A||}$.

Задача 2.

Найти $\nabla f, \nabla^2 f$, где:

- $f(x) = -e^{-x^T x}$
- $f(x) = \dfrac{1}{1+x^T x}$
- $f(x) = \log \sum\limits_{i=1}^{m} \exp\left(a_i^T x + b_i\right), a_i \in \mathbb{R}^n, b_i \in \mathbb{R}$

Решение:

1) Пусть $x = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}$, тогда:

$$-x^T x = -\sum_{i=1}^{n} x_i^2 \Longrightarrow -e^{-x^T x} = -e^{-\left(x_1^2+\ldots+x_n^2\right)}$$

$$\frac{\partial f(x)}{\partial x_k} = 2x_k e^{-\left(x_1^2+\ldots+x_n^2\right)} \Longrightarrow \nabla f(x) = 2xe^{-x^T x}$$

$$\frac{\partial^2 f(x)}{\partial x_k^2} = 2e^{-\left(x_1^2+\ldots+x_n^2\right)} - 4x_k^2 e^{-\left(x_1^2+\ldots+x_n^2\right)} = 2\left(1-2x_k^2\right)e^{-\left(x_1^2+\ldots+x_n^2\right)}$$

Т.к. частные производные непрерывны, порядок второго дифференцирования не имеет значения:

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}\left(2x_i e^{-\left(x_1^2+\ldots+x_n^2\right)}\right) = -4x_i x_j e^{-\left(x_1^2+\ldots+x_n^2\right)}$$

$$\nabla^2 f(x) = \begin{bmatrix} \dfrac{\partial^2 f(x)}{\partial x_1^2} & \dots & \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(x)}{\partial x_n \partial x_1} & \dots & \dfrac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix} =$$

$$= \begin{bmatrix} 2\left(1-2x_1^2\right)e^{-\left(x_1^2+\ldots+x_n^2\right)} & \dots & -4x_1 x_n e^{-\left(x_1^2+\ldots+x_n^2\right)} \\ \vdots & \ddots & \vdots \\ -4x_n x_1 e^{-\left(x_1^2+\ldots+x_n^2\right)} & \dots & 2\left(1-2x_n^2\right)e^{-\left(x_1^2+\ldots+x_n^2\right)} \end{bmatrix} =$$

$$= \left(-4xx^T + 2E\right)e^{-\left(x_1^2+\ldots+x_n^2\right)}$$

2)

$$\frac{\partial f(x)}{\partial x_k} = -\frac{\frac{\partial\left(x^T x\right)}{\partial x_k}}{\left(1+x^T x\right)^2} = -\frac{2x_k}{\left(1+x^T x\right)^2} \Longrightarrow \nabla f(x) = -\frac{2x}{\left(1+x^T x\right)^2}$$

$$\frac{\partial^2 f(x)}{\partial x_k^2} = -\frac{2}{\left(1+x^T x\right)^2} + \frac{8x_k^2}{\left(1+x^T x\right)^3} = \frac{8x_k^2 - 2x^T x - 2}{\left(1+x^T x\right)^3}$$

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{8x_i x_j}{\left(1+x^T x\right)^3}$$

$$\nabla^2 f(x) = \begin{bmatrix} \dfrac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix} =$$

$$= \begin{bmatrix} \dfrac{8x_1^2 - 2x^T x - 2}{\left(1+x^T x\right)^3} & \cdots & \dfrac{8x_1 x_n}{\left(1+x^T x\right)^3} \\ \vdots & \ddots & \vdots \\ \dfrac{8x_n x_1}{\left(1+x^T x\right)^3} & \cdots & \dfrac{8x_n^2 - 2x^T x - 2}{\left(1+x^T x\right)^3} \end{bmatrix} = \frac{8xx^T - 2\left(x^T x + 1\right)E}{\left(1+x^T x\right)^3}$$

3)

$$df = d\left(\log\left(\sum_{i=1}^m e^{\langle a_i, x\rangle + b_i}\right)\right) = \frac{\sum_{i=1}^m d\left(e^{\langle a_i, x\rangle + b_i}\right)}{\sum_{i=1}^m e^{\langle a_i, x\rangle + b_i}} = \frac{\sum_{i=1}^m \left(e^{\langle a_i, x\rangle + b_i}\right) d(\langle a_i, x\rangle)}{\sum_{i=1}^m e^{\langle a_i, x\rangle + b_i}} =$$

$$= \frac{\sum_{i=1}^m \left(e^{\langle a_i, x\rangle + b_i}\right) \langle a_i, dx\rangle}{\sum_{i=1}^m e^{\langle a_i, x\rangle + b_i}} = \left\langle \frac{\sum_{i=1}^m \left(e^{\langle a_i, x\rangle + b_i}\right) a_i}{\sum_{i=1}^m e^{\langle a_i, x\rangle + b_i}}, dx \right\rangle$$

$$\nabla f = \frac{\sum_{i=1}^m \left(e^{\langle a_i, x\rangle + b_i}\right) a_i}{\sum_{i=1}^m e^{\langle a_i, x\rangle + b_i}}$$

$$d\left(\left\langle \frac{\sum_{i=1}^m \left(e^{\langle a_i, x\rangle + b_i}\right) a_i}{\sum_{i=1}^m e^{\langle a_i, x\rangle + b_i}}, dx_1 \right\rangle\right) = \left\langle d\left(\frac{\sum_{i=1}^m \left(e^{\langle a_i, x\rangle + b_i}\right) a_i}{\sum_{i=1}^m e^{\langle a_i, x\rangle + b_i}}\right), dx_1 \right\rangle =$$

$$= \left\langle \frac{\sum_{j=1}^m e^{\langle a_j, x\rangle + b_j} d\left(\sum_{i=1}^m \left(e^{\langle a_i, x\rangle + b_i}\right) a_i\right) - d\left(\sum_{j=1}^m e^{\langle a_j, x\rangle + b_j}\right) \sum_{i=1}^m \left(e^{\langle a_i, x\rangle + b_i}\right) a_i}{\left(\sum_{i=1}^m e^{\langle a_i, x\rangle + b_i}\right)^2}, dx_1 \right\rangle =$$

$$= \left\langle \frac{\sum_{j=1}^m e^{\langle a_j, x\rangle + b_j} \sum_{i=1}^m a_i \left(e^{\langle a_i, x\rangle + b_i}\right) d(\langle a_i, x\rangle)}{\left(\sum_{i=1}^m e^{\langle a_i, x\rangle + b_i}\right)^2}, dx_1 \right\rangle -$$

$$- \left\langle \frac{\sum_{j=1}^m e^{\langle a_j, x\rangle + b_j} d(\langle a_j, x\rangle) \sum_{i=1}^m a_i \left(e^{\langle a_i, x\rangle + b_i}\right)}{\left(\sum_{i=1}^m e^{\langle a_i, x\rangle + b_i}\right)^2}, dx_1 \right\rangle =$$

$$= \left\langle \frac{\sum_{j=1}^m e^{\langle a_j, x\rangle + b_j} \sum_{i=1}^m a_i \left(e^{\langle a_i, x\rangle + b_i}\right) a_i^T dx_2}{\left(\sum_{i=1}^m e^{\langle a_i, x\rangle + b_i}\right)^2}, dx_1 \right\rangle -$$

$$-\left\langle \frac{\sum_{j=1}^{m} e^{\langle a_j,x\rangle+b_j} a_j^T dx_2 \sum_{i=1}^{m} a_i \left(e^{\langle a_i,x\rangle+b_i}\right)}{\left(\sum_{i=1}^{m} e^{\langle a_i,x\rangle+b_i}\right)^2}, dx_1 \right\rangle =$$

$$\left\langle \frac{\sum_{j=1}^{m} e^{\langle a_j,x\rangle+b_j} \sum_{i=1}^{m} \left(e^{\langle a_i,x\rangle+b_i}\right) a_i a_i^T - \sum_{i=1}^{m} a_i \left(e^{\langle a_i,x\rangle+b_i}\right) \sum_{j=1}^{m} e^{\langle a_j,x\rangle+b_j} a_j^T}{\left(\sum_{i=1}^{m} e^{\langle a_i,x\rangle+b_i}\right)^2} dx_2, dx_1 \right\rangle$$

$$\nabla^2 f = \frac{\sum_{j=1}^{m} e^{\langle a_j,x\rangle+b_j} \sum_{i=1}^{m} \left(e^{\langle a_i,x\rangle+b_i}\right) a_i a_i^T - \sum_{j=1}^{m} e^{\langle a_j,x\rangle+b_j} a_j \sum_{i=1}^{m} a_i^T \left(e^{\langle a_i,x\rangle+b_i}\right)}{\left(\sum_{i=1}^{m} e^{\langle a_i,x\rangle+b_i}\right)^2}$$

Ответ:

1)

$$\nabla f(x) = 2x e^{-\left(x_1^2+\ldots+x_n^2\right)}$$

$$\nabla^2 f(x) = \left(-4xx^T + 2E\right) e^{-\left(x_1^2+\ldots+x_n^2\right)}$$

2)

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

$$\nabla^2 f(x) = \frac{8xx^T - 2\left(x^T x + 1\right) E}{\left(1+x^T x\right)^3}$$

3)

$$\nabla f = \frac{\sum_{i=1}^{m} \left(e^{\langle a_i,x\rangle+b_i}\right) a_i}{\sum_{i=1}^{m} e^{\langle a_i,x\rangle+b_i}}$$

$$\nabla^2 f = \frac{\sum_{j=1}^{m} e^{\langle a_j,x\rangle+b_j} \sum_{i=1}^{m} \left(e^{\langle a_i,x\rangle+b_i}\right) a_i a_i^T - \sum_{j=1}^{m} e^{\langle a_j,x\rangle+b_j} a_j \sum_{i=1}^{m} a_i^T \left(e^{\langle a_i,x\rangle+b_i}\right)}{\left(\sum_{i=1}^{m} e^{\langle a_i,x\rangle+b_i}\right)^2}$$

Задача 3.

Найти $\nabla f(X)$, где $f(X) = X^{-1}, X \in S_+^n$ -- положительно определенная симметричная матрица.

Решение:

$$\frac{\partial f}{\partial X} : T = \lim_{\alpha \to 0} \frac{f(X+\alpha T) - f(X)}{\alpha} = \lim_{\alpha \to 0} \frac{(X+\alpha T)^{-1} - X^{-1}}{\alpha} =$$

$$= \lim_{\alpha \to 0} \frac{\left(E + X^{-1}\alpha T\right)^{-1} X^{-1} - X^{-1}}{\alpha} = \lim_{\alpha \to 0} \frac{\left(\sum_{k=0}^{\infty} (-\alpha)^k \left(X^{-1}T\right)^k\right) X^{-1} - X^{-1}}{\alpha}$$

Заметим, что:

$$\sum_{k=0}^{\infty} (-\alpha)^k \left(X^{-1}T\right)^k = E - \alpha X^{-1}T + \alpha^2 \left(X^{-1}T\right)^2 + \ldots = E - \alpha X^{-1}T + o(\alpha)$$

Тогда:

$$\lim_{\alpha \to 0} \frac{\left(\sum_{k=0}^{\infty} (-\alpha)^k \left(X^{-1}T\right)^k\right) X^{-1} - X^{-1}}{\alpha} = \lim_{\alpha \to 0} \frac{\left(E - \alpha X^{-1}T + o(\alpha)\right) X^{-1} - X^{-1}}{\alpha} =$$

$$= \lim_{\alpha \to 0} \frac{X^{-1} - \alpha X^{-1}TX^{-1} + o(\alpha) - X^{-1}}{\alpha} = \lim_{\alpha \to 0} \frac{-\alpha X^{-1}TX^{-1} + o(\alpha)}{\alpha} = -X^{-1}TX^{-1}$$

Итак, **ответ:**

$$\frac{\partial f}{\partial X} : T = -X^{-1}TX^{-1}$$