

Методы оптимизации. Векторное дифференцирование

Александр Катруца

Московский физико-технический институт

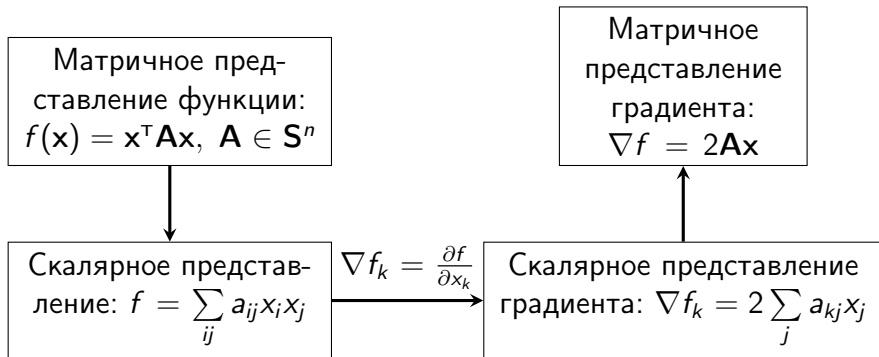
12 февраля 2019 г.

Основные определения

Более подробно смотрите [здесь](#). Пусть $f : D \rightarrow E$,
производная $\frac{\partial f}{\partial x} \in G$:

D	E	G	Название
\mathbb{R}	\mathbb{R}	\mathbb{R}	Производная, $f'(x)$
\mathbb{R}^n	\mathbb{R}	\mathbb{R}^n	Градиент, $\frac{\partial f}{\partial x_i}$
\mathbb{R}^n	\mathbb{R}^m	$\mathbb{R}^{n \times m}$	Матрица Якоби, $\frac{\partial f_i}{\partial x_j}$
$\mathbb{R}^{m \times n}$	\mathbb{R}	$\mathbb{R}^{m \times n}$	$\frac{\partial f}{\partial x_{ij}}$

Также квадратная $n \times n$ матрица вторых производных
 $\mathbf{H} = [h_{ij}]$ в случае $f : \mathbb{R}^n \rightarrow \mathbb{R}$ называется гессиан и равна
$$h_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$



Примеры

1. Линейная функция: $f(\mathbf{x}) = \mathbf{c}^T \mathbf{x}$
2. Квадратичная форма: $f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x}$
3. Квадрат ℓ_2 нормы разности: $f(\mathbf{x}) = \|\mathbf{A} \mathbf{x} - \mathbf{b}\|_2^2$
4. Детерминант: $f(\mathbf{X}) = \det \mathbf{X}$
5. След: $f(\mathbf{X}) = \text{Tr}(\mathbf{A} \mathbf{X} \mathbf{B})$
6. $f(\mathbf{x}) = (\mathbf{x} - \mathbf{A} \mathbf{s})^T \mathbf{W} (\mathbf{x} - \mathbf{A} \mathbf{s})$
7. $f(\mathbf{A}) = (\mathbf{x} - \mathbf{A} \mathbf{s})^T \mathbf{W} (\mathbf{x} - \mathbf{A} \mathbf{s})$
8. $f(\mathbf{s}) = (\mathbf{x} - \mathbf{A} \mathbf{s})^T \mathbf{W} (\mathbf{x} - \mathbf{A} \mathbf{s})$

Сложная функция

Пусть $f(\mathbf{x}) = g(u(\mathbf{x}))$, тогда $\nabla f(\mathbf{x}) = \frac{\partial g}{\partial u} \frac{\partial u}{\partial \mathbf{x}}$

Важно смотреть на размерности и понимать как записывать $\frac{\partial g}{\partial u}$.

Примеры:

1. ℓ_2 норма вектора: $f(\mathbf{x}) = \|\mathbf{x}\|_2$
2. Экспонента: $f(\mathbf{x}) = -e^{-\mathbf{x}^T \mathbf{x}}$

- Производная по скаляру
- Производная по вектору
- Производная по матрице
- Производная сложной функции