Методы оптимизации. Семинар 5. Векторное дифференцирование.

Александр Катруца

Московский физико-технический институт,
Факультет Управления и Прикладной Математики

3 октября 2016 г.

Напоминание

- Сопряжённые множества
- Свойства сопряжённых множеств
- Лемма Фаркаша

Основные определения

Более подробно смотрите здесь. Пусть $f : D \to E$, 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, $\frac{\partial f}{\partial x_i}$
$\mathbb{R}^n$	$\mathbb{R}^m$	$\mathbb{R}^{n\times m}$	Якобиан, $\frac{\partial f_i}{\partial x_j}$
$\mathbb{R}^{m\times n}$	$\mathbb{R}$	$\mathbb{R}^{m\times n}$	$\frac{\partial f}{\partial x_{ij}}$

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

Примеры

1. Линейная функция: $f(\mathbf{x}) = \mathbf{c}^\top \mathbf{x}$
2. Квадратичная форма: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x}$
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: $f(\mathbf{X}) = \mathrm{Tr}(\mathbf{A}\mathbf{X}\mathbf{B})$
6. $f(\mathbf{x}) = (\mathbf{x} - \mathbf{A}\mathbf{s})^\top \mathbf{W}(\mathbf{x} - \mathbf{A}\mathbf{s})$
7. $f(\mathbf{A}) = (\mathbf{x} - \mathbf{A}\mathbf{s})^\top \mathbf{W}(\mathbf{x} - \mathbf{A}\mathbf{s})$
8. $f(\mathbf{s}) = (\mathbf{x} - \mathbf{A}\mathbf{s})^\top \mathbf{W}(\mathbf{x} - \mathbf{A}\mathbf{s})$

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. Билинейная форма: $f(\mathbf{x}) = u^\mathsf{T}(\mathbf{x})\mathbf{R}v(\mathbf{x})$, $\mathbf{R} \in \mathbb{R}^{m\times n}$
3. Экспонента: $f(\mathbf{x}) = -e^{-\mathbf{x}^\mathsf{T}\mathbf{x}}$

Резюме

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