Методы оптимизации. Семинар 2. Векторное дифференцирование

Александр Катруца

Московский физико-технический институт
Факультет Инноваций и Высоких Технологий

11 сентября 2018 г.

Напоминание

- Зачем нужно решать оптимизационный задачи?
- Постановки задач оптимизации
- Выпуклые множества и конусы

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, $\frac{\partial f}{\partial x_i}$
$\mathbb{R}^n$	$\mathbb{R}^m$	$\mathbb{R}^{n\times m}$	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 $f : \mathbb{R}^n \to \mathbb{R}$ называется гессиан и равна $h_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$.

Основная техника

Матричное представление функции: $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x}, \ \mathbf{A} \in \mathbf{S}^n$

Скалярное представление: $f = \sum\limits_{ij} a_{ij} x_i x_j$

$$\nabla f_i = \frac{\partial f}{\partial x_i}$$

Скалярное представление градиента: $\nabla f_i = 2 \sum\limits_{j} a_{ij} x_j$

Матричное представление градиента: $\nabla f = 2\mathbf{A}\mathbf{x}$

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. Экспонента: $f(\mathbf{x}) = -e^{-\mathbf{x}^\top \mathbf{x}}$

Резюме

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