

# Optimization Methods.

## Seminar 5. Introduction to Matrix calculus.

Alexandr Katrutsa

Moscow Institute of Physics and Technology  
Department of Control and Applied Mathematics

October 3, 2016

# Reminder

- Conjugate sets
- Properties of conjugate sets
- Farkas' lemma

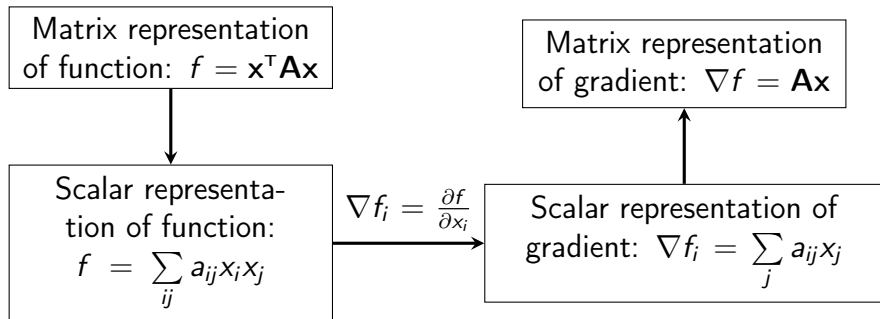
# Basic definitions

More details see [here](#). Let  $f : D \rightarrow E$  be some function, then its derivative  $\frac{\partial f}{\partial x} \in G$ :

$D$	$E$	$G$	Name
$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	Derivative, $f'(x)$
$\mathbb{R}^n$	$\mathbb{R}$	$\mathbb{R}^n$	Gradient, $\frac{\partial f}{\partial x_j}$
$\mathbb{R}^n$	$\mathbb{R}^m$	$\mathbb{R}^{n \times m}$	Jacobian, $\frac{\partial f_i}{\partial x_j}$
$\mathbb{R}^{m \times n}$	$\mathbb{R}$	$\mathbb{R}^{m \times n}$	$\frac{\partial f}{\partial x_{ij}}$

Also square matrix  $n \times n$  of the second derivatives  $\mathbf{H} = [h_{ij}]$  in case of  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is called hessian and its elements equal 
$$h_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

# The main technique



# Examples

- ① Linear function:  $f(\mathbf{x}) = \mathbf{c}^\top \mathbf{x}$
- ② Quadratic form:  $f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^\top \mathbf{A} \mathbf{x} + \mathbf{b}^\top \mathbf{x}$
- ③  $\ell_2$  norm of difference squared:  $f(\mathbf{x}) = \|\mathbf{A} \mathbf{x} - \mathbf{b}\|_2^2$
- ④ Determinant:  $f(\mathbf{X}) = \det \mathbf{X}$
- ⑤ Trace:  $f(\mathbf{X}) = \text{Tr}(\mathbf{A} \mathbf{X} \mathbf{B})$
- 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  $f(\mathbf{s}) = (\mathbf{x} - \mathbf{A} \mathbf{s})^\top \mathbf{W} (\mathbf{x} - \mathbf{A} \mathbf{s})$

# Function composition

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

Check dimensions and understand how to write  $\frac{\partial g}{\partial u}$ .

Examples:

- 1  $\ell_2$  norm of vector:  $f(\mathbf{x}) = \|\mathbf{x}\|_2$
- 2 Bilinear form:  $f(\mathbf{x}) = \mathbf{u}^\top(\mathbf{x})\mathbf{R}\mathbf{v}(\mathbf{x})$ ,  $\mathbf{R} \in \mathbb{R}^{m \times n}$
- 3 Exponent:  $f(\mathbf{x}) = -e^{-\mathbf{x}^\top \mathbf{x}}$

# Recap

- Derivative by scalar
- Derivative by vector
- Derivative by matrix
- Derivative of function composition