Norwegian University of Science and Technology
Department of Engineering Cybernetics

Exercise 0
Matrix Calculus

Problem 1 (25 %) Definitions

a What is the definition of the gradient of a function?

b What is the definition of the Jacobian of a function?

c Let $f(\mathbf{x})$ be a scalar and $\mathbf{x} \in \mathbb{R}^n$. What will the size of the gradient be?

d Let $\mathbf{f}(\mathbf{x})$ be a column vector of length m and $\mathbf{x} \in \mathbb{R}^n$. What will the size of the Jacobian be?

Problem 2 (25 %) Linear

Let $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$, where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

a Use the definition and calculate $\dfrac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}}$. After calculating it, simplify it into a matrix form. Is this the Jacobian or the gradient of $\mathbf{f}(\mathbf{x})$?

b Can you, without doing any calculations, find $\dfrac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}}$ when $\mathbf{x}$ is a column vector of length n, and $\mathbf{A}$ is a matrix of dimension $m \times n$ (i.e., m rows and n columns)?

Problem 3 (25 %) Nonlinear

Let $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T\mathbf{G}\mathbf{y}$, where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.$$

a What is the dimension of $f(\mathbf{x}, \mathbf{y})$? Is $\nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{y})$ equal to $\dfrac{\partial f(\mathbf{x}, \mathbf{y})}{\partial \mathbf{x}}$ (no calculations are needed)?

b Use the definition and calculate $\nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{y})$. Then write the answer in matrix form.

c Use the definition and calculate $\nabla_{\mathbf{y}} f(\mathbf{x}, \mathbf{y})$. Then write the answer in matrix form.

d Let $f(\mathbf{x}) = \mathbf{x}^T\mathbf{H}\mathbf{x}$, where $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{H} \in \mathbb{R}^{n \times n}$. Find $\nabla f(\mathbf{x})$ using the results from the previous exercises. What will the answer be if $\mathbf{H}$ is symmetric?

Problem 4 (25 %) Common case

This problem is supposed to be solved using matrices only (i.e., the elements of the matrices are irrelevant). During this course you will encounter (1), **many** times. $\mathcal{L}(\dots)$ is a scalar, $\mathbf{G}$ is symmetric and $\mathbf{x}$ is of length n. The remaining vectors and matrices are of appropriate sizes such that the output of each term is a scalar. If you feel there is a need of specifying the sizes to solve the problem, justify your reasoning and set up some variable sizes (i.e., use variables (m,n,p,q, ...) instead of fixed numbers).

I would recommend you to try to utilize what you have learnt in the previous exercises before looking at the "Matrix Calculus"-note. Use this as a test to see if you have understood the rules.

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbf{x}^T\mathbf{G}\mathbf{x} + \boldsymbol{\lambda}^T(\mathbf{C}\mathbf{x} - \mathbf{d}) + \boldsymbol{\mu}^T(\mathbf{E}\mathbf{x} - \mathbf{h}) \tag{1}$$

a Find $\nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$.

b Find $\nabla_{\boldsymbol{\mu}}\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$.

c Find $\nabla_{\boldsymbol{\lambda}}\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$.