NTNU

Kunnskap for en bedre verden

TTK4135 - Optimization and Control

Exercise #0

Author:
Sondre Pedersen

January 11, 2025

Problem 1: Definitions

(a) What is the definition of the gradient of a scalar function $f\colon \mathbb{R}^n \to \mathbb{R}$? Is it a row or column vector?

(b) What is the definition of the Jacobian of a function $f\colon \mathbb{R}^n \to \mathbb{R}^m$? If m = 1, what is the difference between the gradient and the Jacobian?

(a)

The definition of the gradient of a scalar function is $\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}^\top$. The result is a (column) vector field where the direction indicates maximum rate of change for *f*.

(b)

$$\mathbf{J}_\mathbf{f} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial x_1} & \cdots & \frac{\partial \mathbf{f}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \nabla^\top f_1 \\ \vdots \\ \nabla^\top f_m \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}.$$

We can see that if $m = 1$ then the Jacobian is a transposed version of the gradient of f.

Problem 2: Linear functions

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

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

(a) Use the definition and calculate $\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}}$. Is this the Jacobian or the gradient of $\mathbf{f}(\mathbf{x})$?

(b) Can you, without doing any calculations, find $\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}}$ 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)?

(a)

$$\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}} = \frac{\partial}{\partial \mathbf{x}} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{\partial}{\partial \mathbf{x}} \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix} = \mathbf{A}.$$

This is the Jacobian of $\mathbf{f}(\mathbf{x})$. It is clearly not the gradient since the result is a matrix.

(b)

Yes, we see from *(a)* that the result will be $\mathbf{A}$.

Problem 3: Nonlinear/quadratic functions

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

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

(a) What is the dimensions of $f(x, y)$? Is $\nabla f(x, y)$ equal to $\frac{\partial f(x,y)}{\partial x}$ (no calculations are needed)?

(b) Use the definition and calcualte $\nabla_x f(x, y)$. Then write the answer in matrix form.

textbf(c) Use the definition and calculate $\nabla_y f(x, y)$. Then write the answer in matrix form.

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

(a)

We can see by the dimensions of the inputs that the ouput will be a scalar value. Informally: $\mathbf{x}^\top \mathbf{G}\mathbf{y} \implies (1 \times 2)(2 \times 3)(3 \times 1) \implies 1$.

(b)

$$
\begin{aligned}
f(\mathbf{x},\mathbf{y}) &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \\
&= \begin{bmatrix} x_1 g_{11} + x_2 g_{21} & x_1 g_{12} + x_2 g_{22} & x_1 g_{13} + x_2 g_{23} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \\
&= y_1(x_1 g_{11} + x_2 g_{21}) + y_2(x_1 g_{12} + x_2 g_{22}) + y_3(x_1 g_{13} + x_2 g_{23})
\end{aligned}
$$

$$
\begin{aligned}
\nabla_{\mathbf{x}} f(\mathbf{x},\mathbf{y}) &= \begin{bmatrix} y_1 g_{11} + y_2 g_{12} + y_3 g_{13} \\ y_1 g_{21} + y_2 g_{22} + y_3 g_{23} \end{bmatrix} \\
&= \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \\
&= \mathbf{G}\mathbf{y}
\end{aligned}
$$

(c)

$$
\nabla_y f(\mathbf{x},\mathbf{y}) = \begin{bmatrix} x_1 g_{11} + x_2 g_{21} \\ x_1 g_{12} + x_2 g_{22} \\ x_1 g_{13} + x_2 g_{23} \end{bmatrix} = \mathbf{G}^\top \mathbf{x}.
$$

(d)

$$
\nabla_x f(\mathbf{x}) = \nabla_x(\mathbf{x}^\top H x) = \mathbf{H}\mathbf{x} + \mathbf{H}^\top \mathbf{x}.
$$

When $\mathbf{H}$ is symmetric:

$$
\nabla_x f(\mathbf{x}) = \mathbf{H}\mathbf{x} + \mathbf{H}^\top \mathbf{x} = \mathbf{H}\mathbf{x} + \mathbf{H}\mathbf{x} = 2\mathbf{H}\mathbf{x}.
$$

Problem 4: A common case: The Lagrangian

Given

$$
\mathcal{L}(x,\lambda,\mu) = x^\top G x + \lambda^\top (Cx - d) + \mu^\top (Ex - h).
$$

(a) Find $\nabla_x \mathcal{L}(x,\lambda,\mu)$.

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

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

(a)

$$
\nabla_x \mathcal{L}(\mathbf{x},\lambda,\mu) = 2\mathbf{G}\mathbf{x} + \mathbf{C}^\top \lambda + \mathbf{E}^\top \mu.
$$

(b)

$$
\nabla_\mu \mathcal{L}(\mathbf{x},\lambda,\mu) = \mathbf{E}\mathbf{x} - \mathbf{h}.
$$

(c)

$$
\nabla_\lambda \mathcal{L}(\mathbf{x},\lambda,\mu) = \mathbf{C}\mathbf{x} - \mathbf{d}.
$$