Mathematical Foundations of Computer Vision

Example Solutions – Assignment 4

Solution of Exercise No. 1

(a) Let $y = Ax$ with $y \in \mathbb{R}^m$, $x \in \mathbb{R}^n$, and where $A = (a_{ij})$, $A \in \mathbb{R}^{m\times n}$, does not depend on x. Prove or disprove: $\frac{\partial y}{\partial x} = A$

We compute $y = Ax$. For $i = 1, \dots, n$ we obtain:

$$y_i = \sum_{j=1}^{n} a_{ij} \cdot x_j.$$

Differentiating yields

$$\frac{\partial y_i}{\partial x_k} = a_{ik} \text{ for } k = 1, \dots, m.$$

Thus

$$\frac{\partial y}{\partial x} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \dots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \dots & \frac{\partial y_m}{x_n} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = A.$$

(b) Let $f = x^\top Ax$ be given where $f \in \mathbb{R}$, $x \in \mathbb{R}^n$, $A = (a_{ij})$, $A \in \mathbb{R}^{n\times n}$.
Compute $\frac{\partial f}{\partial x}$ for (i) A not symmetric, and for (ii) A symmetric.

We reformulate f:

$$f = x^T Ax = x^T \cdot \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj}x_j \end{pmatrix} = \sum_{k=1}^{n} \left(x_k \cdot \sum_{j=1}^{n} a_{kj}x_j \right).$$

Then for $i = 1, \dots, n$:

$$\frac{\partial f}{\partial x_i} = \frac{\partial}{\partial x_i} \left(\sum_{k=1}^{n} \left(x_k \cdot \sum_{j=1}^{n} a_{kj}x_j \right) \right) = \sum_{j=1}^{n} a_{ij}x_j + \sum_{j=1}^{n} a_{ji}x_j = (Ax)_i + (A^T x)_i = (A + A^T)_i x$$

Thus:

(i) In general: $\frac{\partial f}{\partial x} = (A + A^T)x$.

(ii) If A is symmetric, we obtain $\frac{\partial f}{\partial x} = 2Ax$.

(c) Let $f(z) = y^\top(z)x(z)$ where $z \in \mathbb{R}^n$, $x(z) \in \mathbb{R}^n$, $y(z) \in \mathbb{R}^n$.
Compute $\frac{\partial f}{\partial z}$.

$$\frac{\partial f}{\partial z} = \frac{\partial}{\partial z}(y^\top(z)x(z)) = \left(\frac{\partial}{\partial z} y^\top(z)\right) x(z) + y^\top(z) \left(\frac{\partial}{\partial z} x(z)\right)$$

(d) Let $\varphi(x) = \|x - v\|_2$, where x, $v \in \mathbb{R}^n$. Compute $\frac{\partial \varphi}{\partial x}$.

We have

$$\varphi(x) = \|x - v\|_2 = \left(\sum_{j=1}^{n} (x_j - v_j)^2 \right)^{\frac{1}{2}}.$$

Then

$$\frac{\partial \varphi}{\partial x_i} = \frac{1}{2} \cdot \left(\sum_{j=1}^{n} (x_j - v_j) \right)^{-\frac{1}{2}} \cdot \frac{\partial}{\partial x_i} \left(\sum_{j=1}^{n} (x_j - v_j)^2 \right) = \frac{1}{2} \cdot \frac{1}{\|x - v\|_2} \cdot 2(x_i - v_i) = \frac{x_i - v_i}{\|x - v\|_2}$$

$$\Rightarrow \frac{\partial \varphi}{\partial x} = \frac{(x - v)^\top}{\|x - v\|_2}$$

since φ is a mapping from $\mathbb{R}^n to \mathbb{R}$, so that the Jacobian must be in $\mathbb{R}^{1 \times n}$.

Solution of Exercise No. 2

(a) Let $B, C \in \mathbb{R}^{n \times n}$ with $B = (b_{ij})$, $b_{ij} = b_{ij}(t)$ and $C = (c_{ij})$, $c_{ij} = c_{ij}(t)$. Let $BC = I$. Compute the equation resulting out of

$$\frac{d}{dt}[BC] = \frac{d}{dt}[I].$$

We compute

$$\begin{aligned}
\left(\frac{d}{dt}[BC] \right)_{ij} &= \frac{d}{dt} \left(\sum_{k=1}^{n} b_{ik}(t) c_{kj}(t) \right) \\
&= \sum_{k=1}^{n} \frac{d}{dt} \left(b_{ik}(t) c_{kj}(t) \right) \\
&= \sum_{k=1}^{n} \left(\frac{d}{dt} b_{ik}(t) \right) \cdot c_{kj}(t) + b_{ik}(t) \cdot \left(\frac{d}{dt} c_{kj}(t) \right) \\
&= \left(\frac{d}{dt}[B] \cdot C \right)_{ij} + \left(B \cdot \frac{d}{dt}[C] \right)_{ij} \\
&= \left(\frac{d}{dt}[B] \cdot C + B \cdot \frac{d}{dt}[C] \right)_{ij}.
\end{aligned}$$

Thus

$$\begin{aligned}
\frac{d}{dt}[BC] = \frac{d}{dt}[I] &\Leftrightarrow \frac{d}{dt}[B] \cdot C + B \cdot \frac{d}{dt}[C] = 0 \\
&\Leftrightarrow \dot{B}(t) \cdot C(t) + B(t) \cdot \dot{C}(t) = 0.
\end{aligned}$$

(b) Let $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ be invertible, with $a_{ij} = a_{ij}(t)$. Compute $\frac{d}{dt} \left[A^{-1} \right]$.

We have

$$0 = \frac{d}{dt}(I) = \frac{d}{dt} \left[A^{-1}(t) A(t) \right] = \frac{d}{dt} \left[A^{-1}(t) \right] \cdot A(t) + A^{-1}(t) \cdot \frac{d}{dt} \left[A(t) \right].$$

Thus

$$\begin{aligned}
&\frac{d}{dt} \left[A^{-1}(t) \right] \cdot A(t) = -A^{-1}(t) \cdot \frac{d}{dt} \left[A(t) \right] \\
\Rightarrow &\frac{d}{dt} \left[A^{-1}(t) \right] = -A^{-1}(t) \cdot \dot{A}(t) \cdot A^{-1}(t).
\end{aligned}$$

Solution of Exercise No. 3

(a) Prove that the following implication holds:
If *A can be made similar to a diagonal matrix Λ,* **then** *A is symmetric.*

If A can be made similar to a diagonal matrix Λ, then an orthogonal matrix Q exists such that $A = Q\Lambda Q^T$. We show that the matrix A is indeed symmetric:

$$A^T = (Q\Lambda Q^T)^T = (Q^T)^T \Lambda^T Q^T = Q\Lambda^T Q^T = Q\Lambda Q^T = A,$$

as a diagonal matrix is always symmetric.

(b) Prove that the following assertion holds: For a symmetric matrix A, the eigenvalues are real.

Let λ be an eigenvalue of the matrix A corresponding to the eigenvector x. We want to show that λ is real, hence we have to show that $\lambda = \overline{\lambda}$ is fulfilled. Multiplying $Ax = \lambda x$ with $\overline{x}^T$ yields

$$Ax = \lambda x \quad \Rightarrow \quad \overline{x}^T Ax = \overline{x}^T \lambda x \quad \Rightarrow \quad \overline{x}^T Ax = \lambda \overline{x}^T x.$$

Further, computing the complex conjugate of $Ax = \lambda x$ results in

$$Ax = \lambda x \quad \Rightarrow \quad \overline{Ax} = \overline{A}\overline{x} = \overline{\lambda}\overline{x} = \overline{\lambda x} \quad \Rightarrow \quad A\overline{x} = \overline{\lambda}\overline{x},$$

since $A \in \mathbb{R}^{n\times n}$. Take the transpose

$$\Rightarrow (A\overline{x})^T = (\overline{\lambda}\overline{x})^T \quad \Rightarrow \quad \overline{x}^T A^T = \overline{x}^T \underbrace{\overline{\lambda}^T}_{\in\mathbb{C}^{1\times 1}} = \overline{\lambda}\overline{x}^T \quad \Rightarrow \quad \overline{x}^T A = \overline{\lambda}\overline{x}^T$$

as $A = A^T$. Multiplication with x yields

$$\Rightarrow \overline{x}^T Ax = \overline{\lambda}\overline{x}^T x.$$

Comparison with equation (1) results in

$$\overline{\lambda}\overline{x}^T x = \lambda \overline{x}^T x.$$

As $\overline{x}^T x \in \mathbb{R}$ and $x \neq 0$, we obtain $\overline{\lambda} = \lambda$ and all eigenvalues are real.

(c) Prove that the following assertion holds: For a symmetric matrix A, the eigenvectors to different eigenvalues are orthogonal.

Let λ_1 be an eigenvalue of the matrix A with corresponding eigenvector x_1, and similiarly let λ_2 be an eigenvalue of A with corresponding eigenvector x_2. Both eigenvalues should be different, thus let $\lambda_1 \neq \lambda_2$. Then

$$Ax_1 = \lambda_1 x_1 \tag{1}$$
$$Ax_2 = \lambda_2 x_2 \tag{2}$$

is fulfilled. We have to show that $x_1 \perp x_2$ holds. Therefore we multiply the first equation with x_2^T and the second equation with x_1^T:

$$x_2^T(Ax_1) = x_2^T(\lambda_1 x_1)$$
$$x_1^T(Ax_2) = x_1^T(\lambda_2 x_2).$$

We subtract the second equation from the first equation:

$$x_2^T(Ax_1) - x_1^T(Ax_2) = x_2^T(\lambda_1 x_1) - x_1^T(\lambda_2 x_2).$$

First we simplify the left-hand-side and make use of the symmetry of A:

$$\begin{aligned}
\underbrace{x_2^T(Ax_1)}_{\in\mathbb{R}} - x_1^T(Ax_2) &= (x_2^T(Ax_1))^T - x_1^T Ax_2 \\
&= (Ax_1)^T((x_2)^T)^T - x_1^T Ax_2 \\
&= x_1^T A^T x_2 - x_1^T Ax_2 \\
&= x_1^T Ax_2 - x_1^T Ax_2 \\
&= 0.
\end{aligned}$$

Afterwards we simplify the right-hand-side:

$$\begin{aligned}
x_2^T(\lambda_1 x_1) - x_1^T(\lambda_2 x_2) &= \lambda_1 \underbrace{(x_2^T x_1)}_{\in\mathbb{R}} - \lambda_2(x_1^T x_2) \\
&= \lambda_1(x_2^T x_1)^T - \lambda_2(x_1^T x_2) \\
&= \lambda_1(x_1^T x_2) - \lambda_2(x_1^T x_2) \\
&= (\lambda_1 - \lambda_2) \cdot \langle x_1, x_2 \angle.
\end{aligned}$$

Since $\lambda_1 \neq \lambda_2$ as we assumed, it must hold $\langle x_1, x_2 \angle = 0$, i.e. the vectors x_1 and x_2 are orthogonal.