Question 3

CS663 (Digital Image Processing) Assignment 4

Atishay Jain (210050026)
Cheshta Damor (210050040)
Kanad Shende (210050078)

210050026@iitb.ac.in
210050040@iitb.ac.in
210050078@iitb.ac.in

Contents

Question 3

PART

I

Problem 1
Consider a set of N vectors $\mathcal{X} = \{\boldsymbol{x_1}, \boldsymbol{x_2}, ..., \boldsymbol{x_N}\}$ each in $\mathbb{R}^d$, with average vector $\bar{\boldsymbol{x}}$. We have seen in class that the direction $\boldsymbol{e}$ such that $\sum_{i=1}^{N} \|\boldsymbol{x_i} - \bar{\boldsymbol{x}} - (\boldsymbol{e} \cdot (\boldsymbol{x_i} - \bar{\boldsymbol{x}}))\boldsymbol{e}\|^2$ is minimized, is obtained by maximizing $\boldsymbol{e}^t\boldsymbol{C}\boldsymbol{e}$, where $\boldsymbol{C}$ is the covariance matrix of the vectors in $\mathcal{X}$. This vector $\boldsymbol{e}$ is the eigenvector of matrix $\boldsymbol{C}$ with the highest eigenvalue. Prove that the direction $\boldsymbol{f}$ perpendicular to $\boldsymbol{e}$ for which $\boldsymbol{f}^t\boldsymbol{C}\boldsymbol{f}$ is maximized, is the eigenvector of $\boldsymbol{C}$ with the second highest eigenvalue. For simplicity, assume that all non-zero eigenvalues of $\boldsymbol{C}$ are distinct and that $\text{rank}(\boldsymbol{C}) > 2$. Extend the derivation to handle the case of a unit vector $\boldsymbol{g}$ which is perpendicular to both $\boldsymbol{e}$ and $\boldsymbol{f}$ which maximizes $\boldsymbol{g}^t\boldsymbol{C}\boldsymbol{g}$. [10 points]

Section 1

Proof(a)

Proof

We aim to maximize the expression $\boldsymbol{f}^t\boldsymbol{C}\boldsymbol{f}$, subject to the constraint that $\boldsymbol{f}$ is orthogonal to $\boldsymbol{e}$. We will use the method of Lagrange multipliers. The Lagrangian for this optimization problem is as follows:

$$\mathcal{L}(f) = f^tCf - \lambda_1(f^tf - 1) - \lambda_2(f^te)$$

Here, λ_1 and λ_2 are Lagrange multipliers introduced to enforce constraints. The constraint $\boldsymbol{f}^t\boldsymbol{f} = 1$ ensures that $\boldsymbol{f}$ is a unit vector, and $\boldsymbol{f}^t\boldsymbol{e} = 0$ enforces that $\boldsymbol{f}$ is perpendicular to $\boldsymbol{e}$.
Now to find the maximum of $\boldsymbol{f}^t\boldsymbol{C}\boldsymbol{f}$, we set the gradient of $\mathcal{L}$ with respect to $\boldsymbol{f}$ equal to zero:

$$\boldsymbol{\nabla}\mathcal{L}(f) = 2Cf - 2\lambda_1 f - 2\lambda_2 e$$

We set this gradient to zero to maximize $\mathcal{L}$.

$$Cf - \lambda_1 f - \lambda_2 e = 0 \tag{1.1}$$

Now, we premultiply both sides by $\boldsymbol{e}$ to remove the $\boldsymbol{e}$ term from the equation:

$$e^tCf - \lambda_1 e^t f - \lambda_2 e^t e = 0$$

since $\boldsymbol{e}^t\boldsymbol{e} = 1$ i.e, $\boldsymbol{e}$ is unit vector, therefore

$$e^tCf - \lambda_1 e^t f - \lambda_2 = 0 \tag{1.2}$$

Remark From equation (1.2) , we can infer that $\lambda_2 = 0$. This can be shown from the following equations:

$$e^tCf - \lambda_1 e^t f - \lambda_2 = 0$$

Since $\boldsymbol{e}^t\boldsymbol{e} = 1$, we can write

$$e^tCee^tf - \lambda_1 e^t f - \lambda_2 = 0$$

We also know that $\boldsymbol{Ce} = \lambda\boldsymbol{e}$ since $\boldsymbol{e}$ is the eigenvector of matrix $\boldsymbol{C}$ with the highest eigenvalue(λ). Thus

$$e^t\lambda ee^tf - \lambda_1 e^t f - \lambda_2 = 0$$

$$\lambda e^tf - \lambda_1 e^t f - \lambda_2 = 0$$

Also, $\boldsymbol{f}$ is perpendicular to $\boldsymbol{e}$, therefore $\boldsymbol{e}^t\boldsymbol{f} = 0$ which gives:

$$\lambda_2 = 0 \tag{1.3}$$

Now using equations (1.1) and (1.3), we can write

$$Cf = \lambda_1 f \tag{1.4}$$

This result (1.4) establishes that $\boldsymbol{f}$ is an eigenvector of $\boldsymbol{C}$ with eigenvalue λ_1.

Moreover, we assumed that all eigenvalues of $\boldsymbol{C}$ are distinct and $\boldsymbol{f}^t\boldsymbol{Cf} = \lambda_1$ (this can be seen by premultiplying equation (1.4) by $\boldsymbol{f}^t$) which means that λ_1 is the second highest eigenvalue.

□

This concludes the formal proof that the direction $\boldsymbol{f}$, which is perpendicular to $\boldsymbol{e}$ and maximizes $\boldsymbol{f}^t\boldsymbol{Cf}$, corresponds to the eigenvector of $\boldsymbol{C}$ with the second highest eigenvalue, provided that all non-zero eigenvalues of $\boldsymbol{C}$ are distinct.

SECTION 2

Proof(b)

PROOF We aim to maximize $\boldsymbol{g}^t\boldsymbol{C}\boldsymbol{g}$ subject to the constraints that $\boldsymbol{g}$ is orthogonal to both $\boldsymbol{e}$ and $\boldsymbol{f}$. We will use the method of Lagrange multipliers. The Lagrangian for this optimization problem is as follows:

$$\mathcal{L}(g) = g^tCg - \lambda_3(g^tg - 1) - \lambda_4(g^te) - \lambda_5(g^tf)$$

Here, λ_3, λ_4 and λ_5 are Lagrange multipliers.

- $\boldsymbol{g}^t\boldsymbol{g} = 1$ ensures that $\boldsymbol{g}$ is a unit vector
- $\boldsymbol{g}^t\boldsymbol{e} = 0$ enforces that $\boldsymbol{g}$ is perpendicular to $\boldsymbol{e}$
- $\boldsymbol{g}^t\boldsymbol{f} = 0$ enforces that $\boldsymbol{g}$ is perpendicular to $\boldsymbol{f}$

Now to find the maximum, we set the gradient of $\mathcal{L}$ with respect to $\boldsymbol{g}$ equal to zero:

$$\boldsymbol{\nabla}\mathcal{L}(g) = Cg - \lambda_3g - \lambda_4e - \lambda_5f = 0 \tag{2.1}$$

Now, we premultiply both sides of the equation by $\boldsymbol{e}$:

$$e^tCg - \lambda_3e^tg - \lambda_4e^te - \lambda_5e^tf = 0$$

Since $\boldsymbol{e}^t\boldsymbol{e} = 1$:

$$e^tCg - \lambda_3e^tg - \lambda_4 - \lambda_5e^tf = 0 \tag{2.2}$$

Remark From equation (2.2) , we can infer that $\lambda_4 = 0$. This can be concluded by the following series of equations:

$$e^tCee^tg - \lambda_3e^tg - \lambda_4 - \lambda_5e^tf = 0$$
$$e^t\lambda ee^tg - \lambda_3e^tg - \lambda_4 - \lambda_5e^tf = 0$$
$$\lambda e^tg - \lambda_3e^tg - \lambda_4 - \lambda_5e^tf = 0$$

Now since $\boldsymbol{g}$, $\boldsymbol{e}$ and $\boldsymbol{f}$ are mutually perpendicular, therefore $\boldsymbol{e}^t\boldsymbol{g} = 0$ and $\boldsymbol{e}^t\boldsymbol{f} = 0$. Hence we get

$$\lambda_4 = 0 \tag{2.3}$$

Similarly, premultipling both sides of the gradient equation by $\boldsymbol{f}$, we get

$$\lambda_1f^tg - \lambda_3f^tg - \lambda_4f^te - \lambda_5 = 0$$

$$\lambda_5 = 0 \tag{2.4}$$

Now using equations (2.1), (2.3) and (2.4), we can write

$$Cg = \lambda_3 g \tag{2.5}$$

This result (2.5) establishes that is an eigenvector of $\boldsymbol{C}$ with eigenvalue λ_1.
Moreover, we assumed that all eigenvalues of $\boldsymbol{C}$ are distinct and $\boldsymbol{g}^t\boldsymbol{C}\boldsymbol{g} = \lambda_3$ (this can be seen by premultiplying equation (2.5) by $\boldsymbol{g}^t$) which means that λ_3 is the third highest eigenvalue. □

This concludes the formal proof that the direction $\boldsymbol{g}$, which is perpendicular to both $\boldsymbol{e}$ and $\boldsymbol{f}$ and maximizes $\boldsymbol{g}^t\boldsymbol{C}\boldsymbol{g}$, corresponds to the eigenvector of $\boldsymbol{C}$ with the third highest eigenvalue, provided that all non-zero eigenvalues of $\boldsymbol{C}$ are distinct.