Aalto University

Linear algebra

Exercise sheet 7 / Model solutions

1. Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian matrix. Show that

   (a) $\|\boldsymbol{x}\|_2^2 = \boldsymbol{x}^*\boldsymbol{x} = \sum_{i=1}^n |\alpha_i|^2$,

   (b) $\boldsymbol{x}^*A\boldsymbol{x} = \sum_{i=1}^n \lambda_i|\alpha_i|^2$,

   (c) $\|A\boldsymbol{x}\|_2^2 = \sum_{i=1}^n \lambda_i^2|\alpha_i|^2$,

   where $\alpha_i \in \mathbb{C}$, for $i = 1,\dots,n$, are the coordinates of the vector $\boldsymbol{x} \in \mathbb{C}^n$ in the orthonormal eigenbasis of the matrix A and $\lambda_i \in \mathbb{R}$, for $i = 1,\dots,n$ are the corresponding eigenvalues (repeated by their algebraic multiplicity). That is, one has $\boldsymbol{x} = \sum_{i=1}^n \alpha_i\boldsymbol{v}_i$, where the vectors $\boldsymbol{v}_i$, for $i = 1,\dots,n$, satisfy $A\boldsymbol{v}_i = \lambda_i\boldsymbol{v}_i$, $\|\boldsymbol{v}_i\|_2 = 1$ and $\boldsymbol{v}_i^*\boldsymbol{v}_j = 0$ if $i \neq j$.

   Solution.

   (a) The first equality holds by definition. For the second equality, write $A = Q\Lambda Q^T$, an eigendecomposition of A. By assumption $\boldsymbol{x} = Q\boldsymbol{\alpha}$. Then,

   $$\boldsymbol{x}^*\boldsymbol{x} = \boldsymbol{\alpha}^*Q^*Q\boldsymbol{\alpha} = \sum_{i=1}^n |\alpha_i|^2.$$

   (b) With the same notation as in the above part,

   $$\boldsymbol{x}^*A\boldsymbol{x} = \boldsymbol{\alpha}^*Q^*AQ\boldsymbol{\alpha} = \boldsymbol{\alpha}^*\Lambda\boldsymbol{\alpha} = \sum_{i=1}^n \lambda_i|\alpha_i|^2.$$

   (c) And again, using the same decomposition,

   $$\|A\boldsymbol{x}\|_2^2 = \boldsymbol{x}^*A^*A\boldsymbol{x} = \boldsymbol{\alpha}^*Q^*A^*AQ\boldsymbol{\alpha} = \boldsymbol{\alpha}^*\Lambda^2\boldsymbol{\alpha} = \sum_{i=1}^n \lambda_i^2|\alpha_i|^2.$$

2. Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian matrix.

   (a) Prove the second half of Theorem 2.2 from the lecture notes:

   $$\lambda_{\max}(A) = \max_{0\neq\boldsymbol{x}\in\mathbb{C}^n} R(A,\boldsymbol{x}),$$

   where $\lambda_{\max}(A) \in \mathbb{R}$ is the largest eigenvalue of the matrix A.

(b) Show that for each $\boldsymbol{x} \in \mathbb{C}^n$ it holds that

$$\lambda_{\min}(A)\|\boldsymbol{x}\|_2^2 \leq \boldsymbol{x}^* A \boldsymbol{x} \leq \lambda_{\max}(A)\|\boldsymbol{x}\|_2^2,$$

where $\lambda_{\min}(A), \lambda_{\max}(A) \in \mathbb{R}$ are the smallest and the largest eigenvalue of A.

Solution.

(a) For simplicity, let $\lambda_n = \lambda_{\max}(A)$ denote the largest eigenvalue of A as in the lecture notes. Suppose that $\{\boldsymbol{q}_i\}_{i=1}^n$ is an orthonormal eigenbasis for A. Then any $\boldsymbol{x} \in \mathbb{C}^n$ admits an expansion $\boldsymbol{x} = \sum_{i=1}^n \alpha_i \boldsymbol{q}_i$, with $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{C}^n$. By Problem 1(b),

$$\boldsymbol{x}^* A \boldsymbol{x} = \sum_{i=1}^n \lambda_i |\alpha_i|^2, \qquad \boldsymbol{x}^* \boldsymbol{x} = \sum_{i=1}^n |\alpha_i|^2.$$

Hence for the Rayleigh quotient we have

$$R(A, \boldsymbol{x}) = \frac{\sum_{i=1}^n \lambda_i |\alpha_i|^2}{\sum_{i=1}^n |\alpha_i|^2} \leq \frac{\lambda_n \sum_{i=1}^n |\alpha_i|^2}{\sum_{i=1}^n |\alpha_i|^2} = \lambda_n.$$

On the other hand when $\alpha_1 = \cdots = \alpha_{n-1} = 0$, and $\alpha_n = 1$ we obtain $R(A, \boldsymbol{x}) = \lambda_n$. This proves the statement.

(b) Denote by λ_1, λ_n for short the minimal and maximal eigenvalue of A, respectively. By the proof in the lecture notes and by item (a) we have that $\lambda_1 \leq R(A, \boldsymbol{x}) \leq \lambda_n$; multiplying each side by $\boldsymbol{x}^* \boldsymbol{x}$ we get

$$\lambda_1 \|\boldsymbol{x}\|_2^2 \leq \boldsymbol{x}^* A \boldsymbol{x} \leq \lambda_n \|\boldsymbol{x}\|_2^2.$$

3. Consider the function $\|\cdot\|_{P^1} : \mathbb{R}^2 \to \mathbb{R}$ defined as follows: for any $\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2$, set

$$\|\boldsymbol{x}\|_{P^1}^2 = \frac{1}{3} x_1^2 + x_1 x_2 + x_2^2.$$

(a) Find a symmetric matrix $A \in \mathbb{R}^{2\times 2}$ such that

$$\|\boldsymbol{x}\|_{P^1}^2 = \boldsymbol{x}^T A \boldsymbol{x}, \qquad \text{for all } \boldsymbol{x} \in \mathbb{R}^2.$$

(b) Show that A above is positive definite (see Definition 2.4 from the lecture notes).

(c) Show that $\|\cdot\|_{P^1} : \mathbb{R}^2 \to \mathbb{R}$ is a norm.

Hints: (b) Use Problem 2(b). (c) Use Lemma 2.3. from the lecture notes.

Solution.

(a) Set

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{and} \quad A = A^T = \begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

This way we get

$$\boldsymbol{x}^T A \boldsymbol{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = a x_1^2 + 2b x_1 x_2 + x_2^2.$$

Comparing $\boldsymbol{x}^T A \boldsymbol{x} = a x_1^2 + 2b x_1 x_2 + x_2^2$ and $\|\boldsymbol{x}\|_{P^1}^2 = \frac{1}{3}x_1^2 + x_1 x_2 + x_2^2$, we find that $a = 1/3$, $2b = 1$ and $c = 1$. That is, we get

$$A = \begin{bmatrix} 1/3 & 1/2 \\ 1/2 & 1 \end{bmatrix}.$$

(b) A matrix A is positive definite by definition if $\boldsymbol{x}^T A \boldsymbol{x} > 0$ for all $\boldsymbol{x} \neq \boldsymbol{0}$. By the previous exercise we get

$$\lambda_{\text{min}}(A)\|\boldsymbol{x}\|_2^2 \leq \boldsymbol{x}^* A \boldsymbol{x} \leq \lambda_{\text{max}}(A)\|\boldsymbol{x}\|_2^2$$

and so in particular if $\lambda_{\text{min}}(A) > 0$, then $\boldsymbol{x}^* A \boldsymbol{x} > 0$. Let us then compute the eigenvalues of A by setting $\det(A - \lambda I) = 0$. We get

$$\begin{vmatrix} \frac{1}{3} - \lambda & \frac{1}{2} \\ \frac{1}{2} & 1 - \lambda \end{vmatrix} = \lambda^2 - \frac{4\lambda}{3} + \frac{1}{12} = 0,$$

and the solutions λ for this are $\frac{1}{6}\left(\sqrt{13} + 4\right)$ and $\frac{1}{6}\left(4 - \sqrt{13}\right)$. The first one is clearly positive, and since $4 - \sqrt{13} \approx 0.4$, also the second one is, hence A is positive definite.

(c) Since in the function

$$f(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^T A \boldsymbol{y}$$

A is symmetric and positive definite, f defines an inner product. From the defintion of inner product it follows in turn that $f(\boldsymbol{x}, \boldsymbol{x})$ alwasy defines a norm, so that $\|\cdot\|_{P^1} = f(\cdot, \cdot)$ is a norm.

4. The eigenvalues of a matrix $A \in \mathbb{C}^{n \times n}$ are usually not computed in practice by finding the zeros of the characteristic polynomial. One way to approximate the eigenvector corresponding to the eigenvalue of a diagonalizable matrix A with largest absolute value is power iteration,

$$\boldsymbol{x}_{i+1} = \frac{A\boldsymbol{x}_i}{\|A\boldsymbol{x}_i\|_2}, \qquad i = 0, 1, 2, \ldots,$$

where $\boldsymbol{x}_0 \in \mathbb{C}^n$ is some initial guess. Approximation for the eigenvalue of A with largest absolute value is computed as

$$\mu_i = R(A, \boldsymbol{x}_i), \qquad i = 0, 1, 2, \ldots \tag{1}$$

The intuition behind power iteration is that repeated multiplication by A turns the vector slowly towards the eigenvector corresponding to the eigenvalue of A with largest absolute value (unless $\boldsymbol{x}_0$ is orthogonal to it).

Let now

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

Choose $\boldsymbol{x}_0 = [1, 0, 0]^T$. Compute an approximation for the eigenvalue of A with largest absolute value $\lambda_{\text{amax}}(A)$ and the corresponding eigenvector by using power iteration. Plot the error $|\lambda_{\text{amax}}(A) - \mu_i|$ as a function of the index $i = 0, 1, \ldots, 10$.

Solution. As the iteration progresses, $\boldsymbol{x}_i$ approaches the eigenvector corresponding to the largest eigenvalue of A, with μ_i approaching the largest eigenvalue of A.

Matlab code:

```
A=[2 1 0;1 2 1; 0 1 2];
x=[1 1 1]'; % starting iteration vector, arbitrarily chosen
m=10; % amount of iterations
myy=zeros(m,1); % mx1 zero vector, here will come the values of myy
error=zeros(m,1); % mx1 zero vector, here will come the error
                  % at the correponding iteration step

for k=1:m
x=(A*x)/norm(A*x); % iteration step
myy(k)=(x'*A*x)/(norm(x))^2; % eigenvalue approximation
error(k)=abs(2 + sqrt(2)-myy(k)); % error
end

myy(m)  % prints the approximation of the largest eigenvalue

% let us also draw the picture
figure
plot(linspace(1,m,m), error)
xlabel('number of iterations')
ylabel('error')
```

(i) As approximation of the the largest eigenvalue of A we get, by the above code, $\mu = 3.4142$.

(ii) We compute manually the largest eigenvalue of A to calculate the error:

$$\begin{aligned}\det(A-\lambda I) &= \det\begin{bmatrix}2-\lambda & 1 & 0\\ 1 & 2-\lambda & 1\\ 0 & 1 & 2-\lambda\end{bmatrix}\\ &= (2-\lambda)^3-(2-\lambda)-(2-\lambda)\\ &= (2-\lambda)(\lambda^2-4\lambda+2)=0\end{aligned}$$

We get the eigenvalus $\lambda_1 = 2$, $\lambda_2 = 2+\sqrt{2}$ ja $\lambda_3 = 2-\sqrt{2}$. Of these the largest is $\lambda_2 = 2+\sqrt{2}$, and it's found in the code in the expression `error(k)`. With the chosen vector $\boldsymbol{x}_0$, the iteration converged pretty quickly.

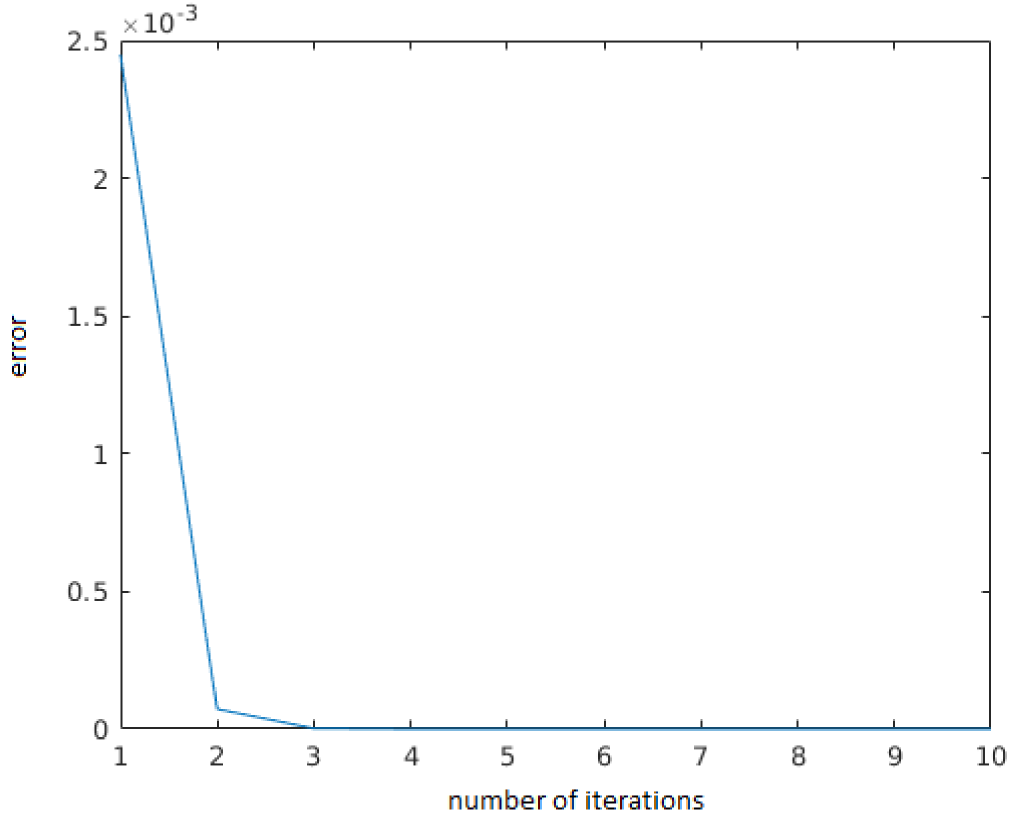

Figure 1: Error