Aalto-yliopisto

Linear algebra

Exercise sheet 3 / Model solutions

1. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be such that

$$f(\boldsymbol{x}) = \left(\boldsymbol{x}^T A \boldsymbol{x}\right)^{1/2} \quad \text{where} \quad A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

The matrix A can be decomposed as $A = U^T \Lambda U$, where

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad \text{and} \quad \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}.$$

(a) Show that $\boldsymbol{y}^T \Lambda \boldsymbol{y} \geq \boldsymbol{y}^T \boldsymbol{y}$ for all $\boldsymbol{y} \in \mathbb{R}^2$ and $\boldsymbol{x}^T A \boldsymbol{x} \geq \boldsymbol{x}^T \boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^2$. (*Hint:* Write $A = U^T \Lambda U$ and set $\boldsymbol{y} = U\boldsymbol{x}$.)

(b) Show that $\boldsymbol{x}^T A \boldsymbol{y} \leq \left(\boldsymbol{x}^T A \boldsymbol{x}\right)^{1/2} \left(\boldsymbol{y}^T A \boldsymbol{y}\right)^{1/2}$. (*Hint:* Write $\boldsymbol{x}^T A \boldsymbol{y} = (\Lambda^{1/2} U \boldsymbol{x})^T (\Lambda^{1/2} U \boldsymbol{y})$ and use the Cauchy–Schwarz inequality, $|\boldsymbol{x}^T \boldsymbol{y}| \leq \|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_2$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$.)

(c) Show that f satisfies the triangle inequality : $f(\boldsymbol{x} + \boldsymbol{y}) \leq f(\boldsymbol{x}) + f(\boldsymbol{y})$. Is f a norm?

Solution.

(a) By direct calculation,

$$\boldsymbol{y}^T \Lambda \boldsymbol{y} = y_1^2 + 3y_2^2 \geq y_1^2 + y_2^2 = \boldsymbol{y}^T \boldsymbol{y}. \tag{1}$$

Using the decomposition of A,

$$\boldsymbol{x}^T A \boldsymbol{x} = \boldsymbol{x}^T U^T \Lambda U \boldsymbol{x} = (U\boldsymbol{x})^T \Lambda (U\boldsymbol{x}).$$

Denoting $\boldsymbol{y} = U\boldsymbol{x}$ and using (1),

$$(U\boldsymbol{x})^T A (U\boldsymbol{x}) = \boldsymbol{y}^T \Lambda \boldsymbol{y} \geq \boldsymbol{y}^T \boldsymbol{y} = \boldsymbol{x}^T U^T U \boldsymbol{x}.$$

Noticing that $U^T U = I$ completes the proof.

(b) Using the decomposition and hint gives

$$\boldsymbol{x}^T A \boldsymbol{y} = \boldsymbol{x}^T U^T \Lambda U \boldsymbol{x} = (\Lambda^{1/2} U \boldsymbol{x})^T (\Lambda^{1/2} U \boldsymbol{y}).$$

Application of the Cauchy–Schwarz inequality to the RHS gives:

$$|(\Lambda^{1/2} U \boldsymbol{x})^T (\Lambda^{1/2} U \boldsymbol{y})| \leq \|\Lambda^{1/2} U \boldsymbol{x}\|_2 \|\Lambda^{1/2} U \boldsymbol{y}\|_2$$

The Euclidian norm is defined as $\|\boldsymbol{x}\|_2 = \left(\boldsymbol{x}^T \boldsymbol{x}\right)^{1/2}$ so that

$$\boldsymbol{x}^T A \boldsymbol{y} \leq \|\Lambda^{1/2} U \boldsymbol{x}\|_2 \|\Lambda^{1/2} U \boldsymbol{y}\|_2 = \left(\boldsymbol{x}^T U^T \Lambda U \boldsymbol{x}\right)^{1/2} \left(\boldsymbol{y}^T U^T \Lambda U \boldsymbol{y}\right)^{1/2},$$

which completes the proof.

(c) By direct calculation and by symmetry of A,

$$(\boldsymbol{x}+\boldsymbol{y})^T A(\boldsymbol{x}+\boldsymbol{y}) = \boldsymbol{x}^T A\boldsymbol{x} + 2\boldsymbol{x}^T A\boldsymbol{y} + \boldsymbol{y}^T A\boldsymbol{y}.$$

Using (b) we estimate this from above as

$$(\boldsymbol{x}+\boldsymbol{y})^T A(\boldsymbol{x}+\boldsymbol{y}) \le \boldsymbol{x}^T A\boldsymbol{x} + 2(\boldsymbol{x}^T A\boldsymbol{x})^{1/2}(\boldsymbol{y}^T A\boldsymbol{y})^{1/2} + \boldsymbol{y}^T A\boldsymbol{y},$$

which can be written as

$$(\boldsymbol{x}+\boldsymbol{y})^T A(\boldsymbol{x}+\boldsymbol{y}) \le \left((\boldsymbol{x}^T A\boldsymbol{x})^{1/2} + (\boldsymbol{y}^T A\boldsymbol{y})^{1/2}\right)^2.$$

Taking a square root from both sides completes the proof that $f(\boldsymbol{x}+\boldsymbol{y}) \le f(\boldsymbol{x}) + f(\boldsymbol{y})$. Lastly, f is indeed a norm, as it also satisfies the other two requirements because $f(\boldsymbol{x}) \ge \|\boldsymbol{x}\|_2$ by item (a) and $f(\alpha\boldsymbol{x}) = \sqrt{\alpha^2\boldsymbol{x}^T A\boldsymbol{x}} = |\alpha| f(\boldsymbol{x})$.

2. Let $\boldsymbol{q}_1 = [0,1,1]^T$, $\boldsymbol{q}_2 = [1,1,-1]^T$ and $\boldsymbol{q}_3 = [2,-1,1]^T$.

   (a) Show that $\boldsymbol{q}_1$, $\boldsymbol{q}_2$ and $\boldsymbol{q}_3$ are mutually orthogonal with respect to the Euclidian inner product.

   (b) Represent the vectors $\boldsymbol{v} = [1,2,3]^T$ and $\boldsymbol{w} = [4,-2,1]^T$ as a linear combination of the vectors $\boldsymbol{q}_1$, $\boldsymbol{q}_2$ and $\boldsymbol{q}_3$. (*Hint:* Write $\boldsymbol{v} = \alpha_1\boldsymbol{q}_1 + \alpha_2\boldsymbol{q}_2 + \alpha_3\boldsymbol{q}_3$. Use the orthogonality to find the coefficients $\alpha_1, \alpha_2, \alpha_3$.)

Solution.

(a) By a direct computation,

$$\boldsymbol{q}_1\cdot\boldsymbol{q}_2 = 0+1-1 = 0, \qquad \boldsymbol{q}_2\cdot\boldsymbol{q}_3 = 2-1-1 = 0, \qquad \boldsymbol{q}_1\cdot\boldsymbol{q}_3 = 0-1+1 = 0.$$

(b) For each coefficient α_i of $\boldsymbol{v}$ in the orthogonal basis $\{\boldsymbol{q}_i\}_{i=1,\dots,3}$ we can use the formula

$$\alpha_i = \frac{\boldsymbol{q}_i\cdot\boldsymbol{v}}{\boldsymbol{q}_i\cdot\boldsymbol{q}_i},$$

from which we get that

$$\alpha_1 = 5/2, \quad \alpha_2 = 0, \quad \alpha_3 = 1/2.$$

With a similar technique one can show that $\boldsymbol{w} = (-1/2)\boldsymbol{q}_1 + (1/3)\boldsymbol{q}_2 + (11/6)\boldsymbol{q}_3$.

3. Let $\boldsymbol{q}_1 = [0,3,2]^T$, $\boldsymbol{q}_2 = [5,1,-1]^T$ and $\boldsymbol{q}_3 = [2,-2,2]^T$ and define the inner product

$$\langle\boldsymbol{x},\boldsymbol{y}\rangle = x_1y_1 + 2x_2y_2 + 3x_3y_3. \tag{2}$$

   (a) Show that $\boldsymbol{q}_1$, $\boldsymbol{q}_2$ and $\boldsymbol{q}_3$ are mutually orthogonal with respect to the inner product defined in (2).

(b) Represent the vectors $\boldsymbol{v} = [1, 2, 3]^T$ and $\boldsymbol{w} = [4, -2, 1]^T$ as a linear combination of the vectors $\boldsymbol{q}_1$, $\boldsymbol{q}_2$ and $\boldsymbol{q}_3$.

Solution.

(a) By direct calculation,

$$\begin{aligned}
\langle \boldsymbol{q}_1, \boldsymbol{q}_2 \rangle &= 0 \times 5 + 2 \times 3 \times 1 + 3 \times 2 \times (-1) = 6 - 6 = 0, \\
\langle \boldsymbol{q}_1, \boldsymbol{q}_3 \rangle &= 0 \times 2 + 2 \times 3 \times (-2) + 3 \times 2 \times (2) = -12 + 12 = 0, \\
\langle \boldsymbol{q}_2, \boldsymbol{q}_3 \rangle &= 5 \times 2 + 2 \times 1 \times (-2) + 3 \times (-1) \times 2 = 10 - 4 - 6 = 0.
\end{aligned}$$

(b) Similarly to the previous exercise, we may write

$$\boldsymbol{v} = \sum_{i=1}^{3} \frac{\langle \boldsymbol{v}, \boldsymbol{q}_j \rangle}{\langle \boldsymbol{q}_j, \boldsymbol{q}_j \rangle} \boldsymbol{q}_j = \sum_{i=1}^{3} \frac{\langle \boldsymbol{v}, \boldsymbol{q}_j \rangle}{\underbrace{\|\boldsymbol{q}_j\|^2}_{\substack{\text{norm induced} \\ \text{by the inn.prod.}}}} \boldsymbol{q}_j.$$

Let us then compute the coefficients in the sum:

$$\begin{aligned}
\frac{\langle \boldsymbol{v}, \boldsymbol{q}_1 \rangle}{\langle \boldsymbol{q}_1, \boldsymbol{q}_1 \rangle} &= \frac{0 + 12 + 18}{0 + 18 + 12} = 1, \\
\frac{\langle \boldsymbol{v}, \boldsymbol{q}_2 \rangle}{\langle \boldsymbol{q}_2, \boldsymbol{q}_2 \rangle} &= \frac{5 + 4 - 9}{25 + 2 + 3} = 0, \\
\frac{\langle \boldsymbol{v}, \boldsymbol{q}_3 \rangle}{\langle \boldsymbol{q}_3, \boldsymbol{q}_3 \rangle} &= \frac{2 - 8 + 18}{4 + 8 + 12} = \frac{1}{2},
\end{aligned}$$

so that

$$\boldsymbol{v} = \boldsymbol{q}_1 + \frac{1}{2} \boldsymbol{q}_3.$$

Similarly for the vector $\boldsymbol{w}$ we get

$$\begin{aligned}
\frac{\langle \boldsymbol{w}, \boldsymbol{q}_1 \rangle}{\langle \boldsymbol{q}_1, \boldsymbol{q}_1 \rangle} &= \frac{0 - 12 + 6}{30} = -\frac{1}{5}, \\
\frac{\langle \boldsymbol{w}, \boldsymbol{q}_2 \rangle}{\langle \boldsymbol{q}_2, \boldsymbol{q}_2 \rangle} &= \frac{20 - 4 - 3}{30} = \frac{13}{30}, \\
\frac{\langle \boldsymbol{w}, \boldsymbol{q}_3 \rangle}{\langle \boldsymbol{q}_3, \boldsymbol{q}_3 \rangle} &= \frac{8 + 8 + 6}{24} = \frac{11}{12},
\end{aligned}$$

so that

$$\boldsymbol{w} = -\frac{1}{5} \boldsymbol{q}_1 + \frac{13}{30} \boldsymbol{q}_2 + \frac{11}{12} \boldsymbol{q}_3.$$

4. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. Show that:

   (a) $\|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\|_2 \leq \|\boldsymbol{x}\|_1$,

   (b) $\|\boldsymbol{x}\|_1 \leq \sqrt{n}\|\boldsymbol{x}\|_2 \leq n\|\boldsymbol{x}\|_\infty$,

   (c) $|\boldsymbol{x} \cdot \boldsymbol{y}| \leq \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_1$.

   (*Hint:* (b) Write $\|\boldsymbol{x}\|_1$ as an inner product between $\boldsymbol{x}$ and a vector containing elements -1 and 1. Apply the Cauchy–Schwarz inequality.)

   Solution.

   (a) First of all

$$\|\boldsymbol{x}\|_\infty = \max_{j=1,\dots,n} |x_j| = \Big(\max_{j=1,\dots,n} |x_j|^2\Big)^{1/2} \leq \left(\sum_{j=1}^n |x_j|^2\right)^{1/2} = \|\boldsymbol{x}\|_2.$$

   On the other hand

$$\|\boldsymbol{x}\|_1^2 = \left(\sum_{j=1}^n |x_j|\right)^2 = \sum_{j=1}^n |x_j|^2 + \underbrace{\sum_{\substack{j=0\\ j\neq k}}^n \sum_{k=1}^n |x_j||x_k|}_{\geq 0} \geq \sum_{j=1}^n |x_j|^2 = \|\boldsymbol{x}\|_2^2$$

   (b) Define the vector $\hat{\boldsymbol{x}} = [\operatorname{sgn}(x_1), \operatorname{sgn}(x_2), \dots, \operatorname{sgn}(x_n)]^T$, where

$$\operatorname{sgn}(t) = \begin{cases} 1, & \text{if } t \geq 0, \\ -1 & \text{if } t < 0. \end{cases}$$

   Clearly

$$\hat{\boldsymbol{x}} \cdot \boldsymbol{x} = \sum_{j=1}^n \operatorname{sgn}(x_j) x_j = \sum_{j=1}^n |x_j|.$$

   Now the Cauchy–Schwarz inequality gives

$$\|\boldsymbol{x}\|_1 = \hat{\boldsymbol{x}}\cdot\boldsymbol{x} \leq \|\hat{\boldsymbol{x}}\|_2\|\boldsymbol{x}\|_2 = \left(\sum_{j=1}^n \operatorname{sgn}(x_j)^2\right)^{1/2} \|\boldsymbol{x}\|_2 = \left(\sum_{j=1}^n 1\right)^{1/2} \|\boldsymbol{x}\|_2 = \sqrt{n}\|\boldsymbol{x}\|_2.$$

   On the other hand

$$\begin{aligned}\|\boldsymbol{x}\|_2 = \left(\sum_{j=1}^n |x_j|\right)^{1/2} &\leq \left(\sum_{j=1}^n \max_{i=1,\dots,n} |x_i|^2\right)^{1/2} = \left(\sum_{j=1}^n \|\boldsymbol{x}\|_\infty^2\right)^{1/2} \\ &= \big(n\|\boldsymbol{x}\|_\infty^2\big)^{1/2} = \sqrt{n}\|\boldsymbol{x}\|_\infty,\end{aligned}$$

   that is, equivalently,

$$\sqrt{n}\|\boldsymbol{x}\|_2 \leq n\|\boldsymbol{x}\|_\infty.$$

(c) We have

$$
\begin{aligned}
|\boldsymbol{x}\cdot\boldsymbol{y}| &= \left|\sum_{j=1}^{n} x_j y_j\right| \leq \sum_{j=1}^{n} |x_j y_j| \\
&= \sum_{j=1}^{n} |x_j||y_j| \leq \sum_{j=1}^{n} (\underbrace{\max_{i=1,\dots,n} |x_i|}_{=\|\boldsymbol{x}\|_\infty})|y_j| \\
&= \|\boldsymbol{x}\|_\infty \sum_{j=1}^{n} |y_j| = \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_1.
\end{aligned}
$$