Aalto University

Linear algebra

Exercise sheet 9 / Model solutions

1. Let $\boldsymbol{F} : \mathbb{R}^2 \to \mathbb{R}^2$ be defined as

$$\boldsymbol{F}(\boldsymbol{x}) = \begin{bmatrix} 1 - x_1 - x_2 \\ x_1 x_2 \end{bmatrix}.$$

Find the equilibrium points of the differential equation system

$$\boldsymbol{x}'(t) = \boldsymbol{F}(\boldsymbol{x}(t)). \tag{1}$$

Linearise the system (1) close to the equilibrium points and compute the eigenvalues of the corresponding coefficient matrices. How do the solution curves starting close to the equilibrium points behave?

Hint: Write the solution of the linearised system using the variable $\boldsymbol{y}(t) = \boldsymbol{x}(t) - \tilde{\boldsymbol{x}}$, where $\tilde{\boldsymbol{x}} \in \mathbb{R}^2$ is the equilibrium point, as $\boldsymbol{y}(t) = \mathrm{e}^{tA}\boldsymbol{y}_0 = V\mathrm{e}^{t\Lambda}V^{-1}\boldsymbol{y}_0$ for suitable $A = V\Lambda V^{-1}$ and deduce how the solution behaves when $t \to \infty$. Note that you do not have to explicitly find V or V^{-1}.

Solution: The system (1) is

$$\boldsymbol{y}'(t) = \begin{bmatrix} y_1' \\ y_2' \end{bmatrix} = \begin{bmatrix} f_1(y_1, y_2) \\ f_2(y_1, y_2) \end{bmatrix} = \begin{bmatrix} 1 - y_1 - y_2 \\ y_1 y_2 \end{bmatrix} = \boldsymbol{F}(y_1, y_2).$$

Let us look for critical points, that is, points where $\boldsymbol{y}'(t) = 0 = \boldsymbol{F}(\boldsymbol{y})$:

$$\begin{cases} 1 - y_1 - y_2 = 0 \\ y_1 y_2 = 0 \end{cases} \Rightarrow \begin{cases} y_1 + y_2 = 1 \\ y_1 = 0 \text{ or } y_2 = 0 \end{cases} \Rightarrow \boldsymbol{y} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ or } \boldsymbol{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Let us linearise the system in these points: close to the point $\boldsymbol{a}$ we have the approximation $\boldsymbol{F}(\boldsymbol{x}) \approx \boldsymbol{F}(\boldsymbol{a}) + D\boldsymbol{F}(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a})$, where

$$D\boldsymbol{F}(y_1, y_2) = \begin{bmatrix} -1 & -1 \\ y_2 & y_1 \end{bmatrix}.$$

Let us categorise these points by using the eigenvalues of J_F. In the point $\boldsymbol{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ we have

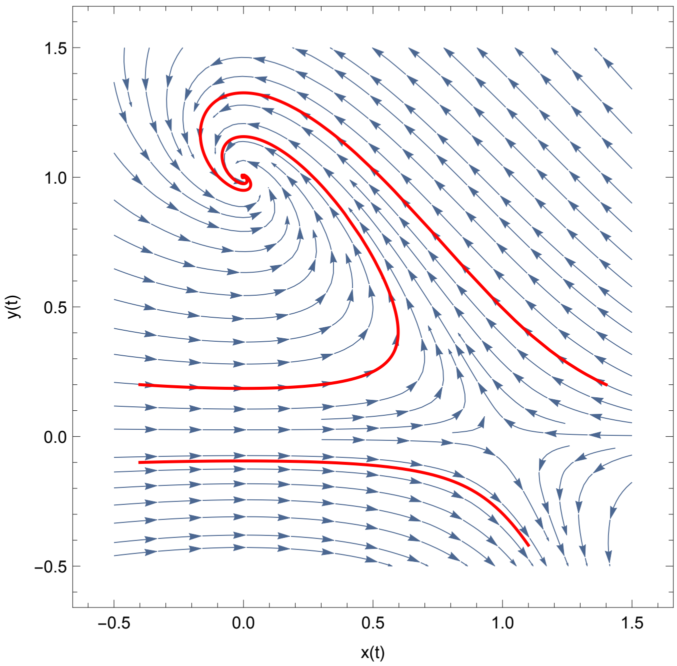
$$D\boldsymbol{F}(0, 1) = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}, \qquad \begin{vmatrix} -1-\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + \lambda + 1 = 0,$$

so that the eigenvalues are $\lambda_{1,2} = -\frac{1}{2} \pm \frac{\sqrt{3}}{2}i$. Since the eigenvalues are complex, but they have a negative real part, the point under consideration is a stable focus.

In the point $\boldsymbol{y} = \begin{bmatrix}1\\0\end{bmatrix}$ we have

$$D\boldsymbol{F}(1,0) = \begin{bmatrix}-1 & -1\\ 0 & 1\end{bmatrix}, \qquad \begin{vmatrix}-1-\lambda & -1\\ 0 & 1-\lambda\end{vmatrix} = \lambda^2 - 1 = 0,$$

so that the eigenvalues are $\lambda_{1,2} = \pm 1$. One eigenvalue is positive and the othe one negative, from which it follows that the equilibrium point $\boldsymbol{y} = \begin{bmatrix}1\\0\end{bmatrix}$ is a saddle point.



The details for the drawing are omitted, left as an exercise.

2. Let $A \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$ and define the function $f : \mathbb{R}^n \to \mathbb{R}$ as

$$f(\boldsymbol{x}) = \|A\boldsymbol{x} - \boldsymbol{b}\|_2^2.$$

Prove by direct calculation that $\nabla f(\boldsymbol{x}) = \boldsymbol{0} \in \mathbb{R}^n$ if and only if $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$.

Hint: You may write $f(\boldsymbol{x}) = (\boldsymbol{x}^T A^T - \boldsymbol{b}^T)(A\boldsymbol{x} - \boldsymbol{b})$ in component form and evaluate the partial derivatives with respect to each coordinate x_ℓ.

Solution: We have that

$$\begin{aligned} f(\boldsymbol{x}) &= (\boldsymbol{x}^T A^T - \boldsymbol{b}^T)(A\boldsymbol{x} - \boldsymbol{b}) \\ &= \boldsymbol{x}^T A^T A\boldsymbol{x} - \boldsymbol{x}^T A^T\boldsymbol{b} - \boldsymbol{b}^T A\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{b} \\ &= \sum_{i,j,k=1}^{n} A_{ki}A_{kj}x_i x_j - 2\sum_{i,j=1}^{n} b_i A_{ij} x_j + \sum_{i=1}^{n} b_i^2. \end{aligned}$$

Hence

$$\begin{aligned}\frac{\partial f}{\partial x_\ell} &= 2\sum_{i,k=1}^{n} A_{ki}A_{k\ell}x_i - 2\sum_{i=1}^{n} b_i A_{i\ell} \\ &= 2\big[(A^TA\boldsymbol{x})_\ell - (A^T\boldsymbol{b})_\ell\big],\end{aligned}$$

from which we find that $\nabla f(\boldsymbol{x})$ has all zero entries if and only if $A^TA\boldsymbol{x} = A^T\boldsymbol{b}$.

3. Let $\boldsymbol{y}, \boldsymbol{b} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$. Consider the minimisation problem

$$\min_{\alpha\in\mathbb{R}} \|\alpha\boldsymbol{y} - \boldsymbol{b}\|^2,$$

where $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ is a norm induced by some inner product $\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, that is, $\|\boldsymbol{z}\|^2 = \langle\boldsymbol{z},\boldsymbol{z}\rangle$ for any $\boldsymbol{z} \in \mathbb{R}^n$.

(a) Show that the minimiser is given by

$$\alpha = \frac{\langle\boldsymbol{y},\boldsymbol{b}\rangle}{\|\boldsymbol{y}\|^2}.$$

(b) Let

$$\boldsymbol{y} = \begin{bmatrix}1\\2\end{bmatrix} \qquad \text{and} \qquad \boldsymbol{b} = \begin{bmatrix}1\\1\end{bmatrix}.$$

Determine α,

(i) when $\|\cdot\|$ is the Euclidian norm,

(ii) when $\|\cdot\|$ is induced by the inner product

$$\langle\boldsymbol{x},\boldsymbol{z}\rangle = \boldsymbol{z}^T\begin{bmatrix}1 & 1\\1 & 2\end{bmatrix}\boldsymbol{x}.$$

Solution:

(a) For any vectors $\boldsymbol{v}$ and $\boldsymbol{u}$ we $\|\boldsymbol{v}+\boldsymbol{u}\|^2 = \|\boldsymbol{v}\|^2 + \|\boldsymbol{u}\|^2 + 2\langle\boldsymbol{v},\boldsymbol{u}\rangle$, so that

$$\|\alpha\boldsymbol{y} - \boldsymbol{b}\|^2 = \alpha^2\|\boldsymbol{y}\|^2 - 2\alpha\langle\alpha,\boldsymbol{b}\rangle + \|\boldsymbol{b}\|^2.$$

The miniminum we are looking for is a root of the derivative (corresponding to the top of the parabola)

$$\frac{d}{d\alpha}\|\alpha\boldsymbol{y} - \boldsymbol{b}\|^2 = -2\langle\boldsymbol{y},\boldsymbol{b}\rangle + 2\|\boldsymbol{y}\|^2\alpha,$$

so that $\alpha = \frac{\langle\boldsymbol{y},\boldsymbol{b}\rangle}{\|\boldsymbol{y}\|^2}$.

(b) (i) From part (a) we get

$$\begin{aligned}\alpha &= (\boldsymbol{y}^T\boldsymbol{y})^{-1}\boldsymbol{y}^T\boldsymbol{b}\\ &= \Big(\begin{bmatrix}1 & 2\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix}\Big)^{-1}\Big(\begin{bmatrix}1 & 2\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}\Big)\\ &= 5^{-1}\times 3 = \frac{3}{5}.\end{aligned}$$

(ii) In this case we have

$$\begin{aligned}\alpha &= (\boldsymbol{y}^TM\boldsymbol{y})^{-1}\boldsymbol{y}^TM\boldsymbol{b}\\ &= \Big(\begin{bmatrix}1 & 2\end{bmatrix}\begin{bmatrix}1 & 1\\1 & 2\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix}\Big)^{-1}\Big(\begin{bmatrix}1 & 2\end{bmatrix}\begin{bmatrix}1 & 1\\1 & 2\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}\Big)\\ &= 13^{-1}\times 8 = \frac{8}{13}.\end{aligned}$$

4. Let $A \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $\langle\cdot,\cdot\rangle : \mathbb{R}^m\times\mathbb{R}^m \to \mathbb{R}$ be some inner product in $\mathbb{R}^m$ and $\|\cdot\|^2 = \langle\cdot,\cdot\rangle$ the corresponding norm. Consider the following least-squares problem: find $\boldsymbol{x} \in \mathbb{R}^n$ such that

$$\|A\boldsymbol{x} - \boldsymbol{b}\|^2 \tag{2}$$

is minimised.

(a) By Lemma 2.3 in Section 1, there exists a symmetric and positive definite matrix $M \in \mathbb{R}^{m\times m}$ such that $\langle\boldsymbol{y},\boldsymbol{z}\rangle = \boldsymbol{z}^TM\boldsymbol{y}$ for all $\boldsymbol{y},\boldsymbol{z} \in \mathbb{R}^m$. Decompose $M = Q\Lambda Q^T$. Show that each eigenvalue of M is positive and find a matrix $L \in \mathbb{R}^{m\times m}$ such that $M = LL^T$.

(b) Use (a) to reformulate the above problem as a least-squares problem posed in the Euclidian norm. That is, find $\tilde{A} \in \mathbb{R}^{m\times n}$ and $\tilde{\boldsymbol{b}} \in \mathbb{R}^m$ such that

$$\|A\boldsymbol{x} - \boldsymbol{b}\|^2 = \|\tilde{A}\boldsymbol{x} - \tilde{\boldsymbol{b}}\|_2^2$$

for all $\boldsymbol{x} \in \mathbb{R}^n$.

Hint: Give the matrix $\tilde{A}$ and vector $\tilde{\boldsymbol{b}}$ as a product of A, $\boldsymbol{b}$ and L. You do not have to define L explicitly, it may remain as a part of the solution.

Solution:

(a) If λ is an eigenvalue of M, then for some nonzero vector $\boldsymbol{x}$ we have $M = \lambda\boldsymbol{x}$. By multiplying on the left by $\boldsymbol{x}^T$ we get

$$\boldsymbol{x}^TM\boldsymbol{x} = \lambda\boldsymbol{x}^T\boldsymbol{x} = \lambda\|\boldsymbol{x}\|_2^2,$$

and $\boldsymbol{x}^TM\boldsymbol{x} > 0$ by assumption, so that $\lambda > 0$. Since M is positive definite, we can diagonalize it as $M = Q\Lambda Q^T$, where the entries of Λ are all positive. We may then write

$$M = Q\sqrt{\Lambda}\sqrt{\Lambda}Q^T = (Q\sqrt{\Lambda})(Q\sqrt{\Lambda})^T = LL^T,$$

where we define $L = Q\sqrt{\Lambda}$. Search "Cholesky decomposition" for additional information.

(b) Let us write $M = LL^T$ and let us compute the least-squares problem with the norm induced by the given inner product:

$$\begin{aligned}\|A\boldsymbol{x} - \boldsymbol{b}\|^2 &= \langle A\boldsymbol{x} - \boldsymbol{b}, A\boldsymbol{x} - \boldsymbol{b}\rangle \\ &= (A\boldsymbol{x} - \boldsymbol{b})^T LL^T (A\boldsymbol{x} - \boldsymbol{b}) \\ &= \boldsymbol{x}^T A^T LL^T A\boldsymbol{x} - 2\boldsymbol{x}^T A^T LL^T \boldsymbol{b} + \boldsymbol{b}^T LL^T \boldsymbol{b}.\end{aligned}$$

Let us differenciate this with respect to $\boldsymbol{x}$ and set it equal to zero:

$$\nabla \left(\boldsymbol{x}^T A^T LL^T A\boldsymbol{x} - 2\boldsymbol{x}^T A^T LL^T \boldsymbol{b} + \boldsymbol{b}^T LL^T \boldsymbol{b}\right) = 0,$$

and this gives

$$A^T LL^T A\boldsymbol{x} = A^T LL^T \boldsymbol{b}.$$

So then we can take $\widetilde{A} = L^T A$ and $\widetilde{\boldsymbol{b}} = L^T \boldsymbol{b}$.