Musterlösung Serie 21

Gram-Schmidt, Orthogonality

1. Let $K = \mathbb{R}, \mathbb{C}$, $A, B \in M_{n\times m}(K)$, $C \in M_{m\times p}(K)$. Prove the following properties of the adjoint matrix:

   (a) $\overline{A+B}^T = \overline{A}^T + \overline{B}^T$;

   (b) For all $\lambda \in K$, $\overline{(\lambda A)}^T = \overline{\lambda}\,\overline{A}^T$;

   (c) $\overline{(\overline{A}^T)}^T = A$;

   (d) $\overline{I_n}^T = I_n$;

   (e) $\overline{(A \cdot C)}^T = \overline{C}^T \cdot \overline{A}^T$.

   Solution: All if these follow from direct computations.

2. Let $K = \mathbb{R}$. On $K[x]_2$, consider the inner product given by

$$\langle p, q\rangle = \int_0^1 p(x)q(x)dx.$$

   (a) Apply the Gram-Schmidt procedure to the basis $1, x, x^2$ to produce an orthonormal basis of $K[x]_2$.

   (b) Find an orthonormal basis of $K[x]_2$ such that the differential operator $p \mapsto p'$ on $K[x]_2$ has an upper triangular matrix with respect to this basis.

   Solution:

   (a) We let $u_1 = 1$ since it is already a vector of norm 1. To apply Gram-Schmidt, we let $v_2 = x - \langle x, 1\rangle$ and $u_2 = \frac{v_2}{||v_2||}$. We compute

$$v_2 = x - \int_0^1 x dx = x - \frac{1}{2}$$

   and

$$||v_2||^2 = \int_0^1 \left(x - \frac{1}{2}\right)^2 dx = \frac{1}{12}.$$

   Hence,

$$u_2 = 2\sqrt{3}\left(x - \frac{1}{2}\right).$$

Similarly, we let $v_3 = x^2 - \langle x^2, u_2\rangle u_2 - \langle x^2, u_1\rangle u_1$ and $u_3 = \frac{v_3}{||v_3||}$. We compute that

$$\begin{aligned} v_3 &= x^2 - \left(\int_0^1 2\sqrt{3}\left(x^3 - \frac{1}{2}x^2\right)dx\right)u_2 - \int_0^1 x^2 dx \\ &= x^2 - \frac{\sqrt{3}}{6}2\sqrt{3}\left(x - \frac{1}{2}\right) - \frac{1}{3} \\ &= x^2 - x + \frac{1}{6}. \end{aligned}$$

Now,

$$\begin{aligned} ||v_3||^2 &= \int_0^1 \left(x^2 - x + \frac{1}{6}\right)^2 dx \\ &= \frac{1}{180}. \end{aligned}$$

Hence

$$u_3 = 6\sqrt{5}\left(x^2 - x + \frac{1}{6}\right).$$

(b) Denote $\mathcal{B} = \{e_1, e_2, e_3\} = \{1, x, x^2\}$, $\mathcal{C} = \{u_1, u_2, u_3\}$ the orthonormal basis found in a), and $D : K[x]_2 \to K[x]_2$, $p \mapsto p'$. We easily compute that

$$[D]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

So, the matrix representation of D with respect to the standard basis $\mathcal{B}$ is already upper triangular, but $\mathcal{B}$ is not orthonormal. However, $\mathcal{C}$ was obtained from $\mathcal{B}$ using the Gram-Schmidt algorithm, so it has the following useful properties:

- $\mathcal{C}$ is orthonormal;
- If for any $i \in \{1, 2, 3\}$, $D(e_i) \in \text{Sp}(e_1, \ldots, e_i)$, the same also holds for $\{u_1, u_2, u_3\}$.

The second property can be seen by analysing the content of the Gram-Schmidt algorithm (try it yourself!), and can be rephrased as follows: if $[D]_{\mathcal{B}}^{\mathcal{B}}$ is upper triangular, then $[D]_{\mathcal{C}}^{\mathcal{C}}$ is upper triangular. Hence $\mathcal{C}$ is a solution.

3. **Minimizing the distance to a subset.** Let K be a field and V be a K-vector space. Suppose that U is a finite-dimensional subspace of V and denote $P_U : V \to U$ the orthogonal projection onto U. Let $v \in V$ and $u \in U$. Show that

$$||v - P_U(v)|| \leqslant ||v - u||.$$

Additionally, prove that the inequality above is an equality if and only if $u = P_U(v)$.

Solution: We have

$$\begin{aligned} ||v - P_U(v)||^2 &\leqslant ||v - P_U(v)||^2 + ||P_U(v) - u||^2 \\ &= ||v - P_U(v) + P_U(v) - u||^2 \\ &= ||v - u||^2 . \end{aligned}$$

Let us justify the above chain of inequality and equality: the first one holds since $||P_U(v) - u|| \geqslant 0$. The second one holds by the Pythagorean theorem. Indeed, by definition of $P_U(v)$, we have $v - P_U(v) \in U^\perp$, and $P_U(v) - u \in U$ since U is a subspace. By taking square roots, we obtain the desired inequality.

The inequality is an equality if and only if equality holds in the first line

$$\begin{aligned} &\Leftrightarrow \quad ||P_U(v) - u||^2 = 0 \\ &\Leftrightarrow \quad P_U(v) = u. \end{aligned}$$

4. Find a polynomial p with real coefficients and degree at most 5 that approximates $\sin(x)$ as well as possible on the interval $[-\pi, \pi]$, in the sense that

$$\int_{-\pi}^{\pi} |\sin(x) - p(x)|^2 \, dx$$

is as small as possible.

Hint. Reformulate the problem in order to use exercise 3.

Solution: Let $V = C([-\pi, \pi], \mathbb{R})$ be the real vector space of real-valued functions on $[-\pi, \pi]$, and consider its subspace U of polynomial functions defined on $[-\pi, \pi]$ of degree at most 5. We endow V with the scalar product

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx, \quad \text{for } f, g \in V.$$

Now, the problem above can be rephrased as follows: find the element of u of U that minimizes $||\sin(x) - u||$, where the norm is induced by the scalar product above. By exercise 3., this quantity will be minimized by the orthogonal projection of $\sin(x)$ to U. To compute it, we will find an orthonormal basis $\mathcal{C} = \{u_0, u_1, \cdots, u_5\}$ of U by applying Gram-Schmidt to the standard basis of U, denoted $\mathcal{B} = \{1, x, x^2, x^3, x^4, x^5\}$, and we will then use the formula

$$P_U(v) = \langle v, u_0 \rangle u_0 + \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_5 \rangle u_5.$$

We make the following observation: since $\sin(x)$ is odd, the integral of its product with any even function on $[-\pi, \pi]$ vanishes. In particular, the integral of its product with even powers of x vanishes. Since odd powers of x are odd functions, we make the same observation.

This greatly simplifies the exercise! It can easily be shown by induction that only basis elements u_i with odd indices will contribute to the computation of further basis elements with odd indices when we apply Gram-Schmidt, and similarly if we replace odd by even. It directly follows that only odd, resp. even, powers of x will appear in u_i's with odd, resp. even, indices. Hence, we only need to compute the coefficients $\langle \sin(x), u_i \rangle$ for odd indices.

So, we compute

$$\begin{aligned} u_1 &= \sqrt{\frac{3}{2\pi^3}}x, \\ u_3 &= \frac{5}{2\pi^{7/2}}\sqrt{\frac{7}{2}}\left(x^3 - \frac{3\pi^2}{5}x\right), \\ u_5 &= \frac{63}{8\pi^{11/2}}\sqrt{\frac{11}{2}}\left(x^5 - \frac{10}{9}\pi^2\left(x^3 - \frac{3\pi^2}{5}x\right) - \frac{3\pi^4}{7}x\right) \end{aligned}$$

and

$$\begin{aligned} P_U(\sin(x)) = &\frac{21}{8\pi^{10}}\big[(33(945 - 105\pi^2 + \pi^4)x^5 - 30\pi^2(1155 - 125\pi^2 + \pi^4)x^3 \\ &+ 5\pi^4(1485 - 153\pi^2 + \pi^4)x\big]. \end{aligned}$$

5. Let $V = C([-1,1], \mathbb{R})$ denote the space of continuous real-valued functions on the interval $[-1,1]$ with inner product

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)dx,$$

for $f, g \in V$. Let $\varphi : V \to \mathbb{R}$ be the linear functional defined by $\varphi(f) = f(0)$. Show that there does not exist $g \in V$ such that

$$\forall f \in V : \varphi(f) = \langle f, g \rangle.$$

Why is this not a counterexample to Satz 6.5.5. in the notes?

Solution: Suppose for a contradiction that such a $g \in V$ exists. Let $f_0 \equiv 1$ on the interval $[-1,1]$. Then

$$1 = f_0(0) = \langle f_0, g \rangle = \int_{-1}^{1} 1 \cdot g(x)dx = \int_{-1}^{1} g(x)dx.$$

Hence there exists some $x_0 \in (-1,1)$ such that $g(x_0) > 0$ and, by continuity, there exists an open interval U_0, containing x_0, such that $g(x) > 0$ on U_0. Shrinking U_0

if needed, we might assume that 0 does not belong to the closure of U_0. Note that, by definition, $h(0) = 0$.

Define $h \in V$ to be a continuous function of $[-1, 1]$ that is strictly positive on a subinterval U_1 of U_0 and vanishes everywhere else, e.g. define a piecewise linear function as done in serie 20, exercise 5.

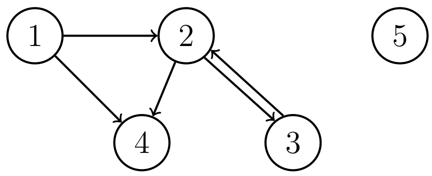
We observe that

$$0 = h(0) = \langle h, g\rangle = \int_{-1}^{1} h(x)g(x)dx = \int_{U_1} h(x)g(x) > 0.$$

This yields a contradiction and proves that such a g does not exist.

6. Let $G = (V, E)$ be a finite directed graph. More precisely, let V be a finite set, and let $E \subseteq \{(v_{\text{init}}, v_{\text{term}}) \mid v_{\text{init}}, v_{\text{term}} \in V \wedge v_{\text{init}} \neq v_{\text{term}}\} \subseteq V \times V$. We think of V as the set of vertices of the graph, and of the pair $(v_{\text{init}}, v_{\text{term}}) \in E$ as the directed edge connecting $v_{\text{init}} \in V$ to $v_{\text{term}} \in V$ (this can be represented by drawing an arrow pointing towards v_{term} on the said edge).

   Example of a directed graph.



   We also define the vector spaces $\mathbb{R}^V = \{f : V \to \mathbb{R}\}$ and $\mathbb{R}^E = \{\varphi : E \to \mathbb{R}\}$, which we equip with the inner products

$$\langle f_1, f_2\rangle_V = \sum_{v \in V} f_1(v)f_2(v), \quad f_1, f_2 \in \mathbb{R}^V$$
$$\langle \varphi_1, \varphi_2\rangle_E = \sum_{e \in E} \varphi_1(e)\varphi_2(e), \quad \varphi_1, \varphi_2 \in \mathbb{R}^E.$$

   Also define $T : \mathbb{R}^V \to \mathbb{R}^E$ as the "combinatorial derivative": for $f \in \mathbb{R}^V$ and $e = (v_{\text{init}}, v_{\text{term}}) \in E$, let

$$T(f)(e) = f(v_{\text{term}}) - f(v_{\text{init}}).$$

   Also define $S : \mathbb{R}^E \to \mathbb{R}^V$ by

$$S(\varphi)(v) = \sum_{\substack{v_{\text{init}} \in V \\ (v_{\text{init}}, v) \in E}} \varphi((v_{\text{init}}, v)) - \sum_{\substack{v_{\text{term}} \in V \\ (v, v_{\text{term}}) \in E}} \varphi((v, v_{\text{term}})).$$

   (a) Show that $T^* = S$ and calculate $T^* \circ T = S \circ T$, which is also called the combinatorial Laplacian of G.

(b) Now simplify the setup by assuming that the graph is undirected, i.e.

$$(v_{\text{init}}, v_{\text{term}}) \in E \Leftrightarrow (v_{\text{term}}, v_{\text{init}}) \in E,$$

and d-regular (for any $v \in V$ there are exactly d vertices $v_{\text{term}} \in V$ with $(v, v_{\text{term}}) \in E$). Show that $T^* \circ T$ admits 0 as an eigenvalue. Explain why the geometric multiplicity of 0 is related to the connectivity of G.

Solution:

(a) For a given edge $e \in E$, let us denote $e^{(1)}$ and $e^{(2)}$ the unique vertices such that $e = (e^{(1)}, e^{(2)})$. For $f \in \mathbb{R}^V$ and $\varphi \in \mathbb{R}^E$, we compute that

$$\begin{aligned}
\langle f, S(\varphi)\rangle_V &= \sum_{v\in V} f(v)S(\varphi)(v) \\
&= \sum_{v\in V} f(v)\left[\sum_{\substack{v_{\text{init}}\in V\\(v_{\text{init}},v)\in E}} \varphi((v_{\text{init}}, v)) - \sum_{\substack{v\in V\\(v,v_{\text{term}})\in E}} \varphi((v, v_{\text{term}}))\right] \\
&= \sum_{v\in V} f(v) \sum_{\substack{v_{\text{init}}\in V\\(v_{\text{init}},v)\in E}} \varphi((v_{\text{init}}, v)) - \sum_{v\in V} f(v) \sum_{\substack{v\in V\\(v,v_{\text{term}})\in E}} \varphi((v, v_{\text{term}})) \\
&= \sum_{e\in E}\sum_{\substack{v\in V\\v=e^{(2)}}} f(v)\varphi(e) - \sum_{e\in E}\sum_{\substack{v\in V\\v=e^{(1)}}} f(v)\varphi(e)
\end{aligned}$$

Since $e^{(1)}$ and $e^{(2)}$ are unique for a fixed $e \in E$, the last line equals

$$\begin{aligned}
\sum_{e\in E} f(e^{(2)})\varphi(e) - \sum_{e\in E} f(e^{(1)})\varphi(e) &= \sum_{e\in E}(f(e^{(2)}) - f(e^{(1)}))\varphi(e) \\
&= \sum_{e\in E} T(f)(e)\varphi(e) \\
&= \langle Tf, \varphi\rangle_E.
\end{aligned}$$

By definition of the adjoint map, this shows that $T^* = S$.
We have

$$\begin{aligned}
S(T(f))(v) &= \sum_{\substack{v_{\text{init}}\in V\\(v_{\text{init}},v)\in E}} T(f)((v_{\text{init}}, v)) - \sum_{\substack{v_{\text{term}}\in V\\(v,v_{\text{term}})\in E}} T(f)((v, v_{\text{term}})) \\
&= \sum_{\substack{v_{\text{init}}\in V\\(v_{\text{init}},v)\in E}} [f(v) - f(v_{\text{init}})] - \sum_{\substack{v_{\text{term}}\in V\\(v,v_{\text{term}})\in E}} [f(v_{\text{term}}) - f(v)] \\
&= E_v f(v) - \sum_{\substack{v_{\text{init}}\in V\\(v_{\text{init}},v)\in E}} f(v_{\text{init}}) - \sum_{\substack{v_{\text{term}}\in V\\(v,v_{\text{term}})\in E}} f(v_{\text{term}}) \\
&= E_v f(v) - \sum_{\substack{w\in V\\w \text{ is a neighbour of } v}} f(w).
\end{aligned}$$

where E_v is the number of edges that touch v.

(b) To solve this question, it is preferable to work with the matrix representation of the combinatorial Laplacian $T^* \circ T \in \mathrm{End}(\mathbb{R}^V)$. Let us label the vertices of V in an arbitrary order as $v_1, \ldots, v_n$. We note that the set characteristic functions

$$\mathcal{B} := \{\mathbf{1}_{x=v_i}(x) \in \mathbb{R}^V \mid i = 1, \ldots, n\}$$

such that

$$\mathbf{1}_{x=v_i}(x) = \begin{cases} 1, & x = v_i \\ 0, & \text{otherwise} \end{cases}$$

forms an orthonormal basis of $(\mathbb{R}^V, \langle \cdot, \cdot \rangle_V)$. Hence a matrix representation of $T^* \circ T$ will be a square matrix of size $n \times n$ whose rows, resp. columns, are indexed by $v_1, \ldots, v_n$, resp. $\mathbf{1}_{x=v_1}, \ldots, \mathbf{1}_{x=v_n}$, in this order.

Let $v \in V$ be arbitrary vertex . We denote $N(v)$ the set of vertices adjacent to v and compute that

$$\begin{aligned}
& (T^* \circ T)(\mathbf{1}_{x=v})(x) \\
= & \sum_{\substack{v_{\text{init}} \in V \\ (v_{\text{init}}, x) \in E}} [\mathbf{1}_{x=v}(x) - \mathbf{1}_{x=v}(v_{\text{init}})] - \sum_{\substack{v_{\text{term}} \in V \\ (x, v_{\text{term}}) \in E}} [\mathbf{1}_{x=v}(v_{\text{term}}) - \mathbf{1}_{x=v}(x)] \\
= & \sum_{w \in N(x)} [\mathbf{1}_{x=v}(x) - \mathbf{1}_{x=v}(w)] - \sum_{w \in N(x)} [\mathbf{1}_{x=v}(w) - \mathbf{1}_{x=v}(x)] \\
= & 2d\mathbf{1}_{x=v}(x) - 2 \sum_{w \in N(x)} \mathbf{1}_{x=v}(x) \\
= & \begin{cases} 2d, & x = v \\ -2, & x \in N(v) \end{cases}
\end{aligned}$$

Note that the factor 2 comes from the fact that we are not working in a directed graph anymore. The elements of $[T^* \circ T]_{\mathcal{B}}^{\mathcal{B}}$ can therefore be described as follows:

$$([T^* \circ T]_{\mathcal{B}}^{\mathcal{B}})_{i,j} = \begin{cases} 2d, & i = j \\ -2, & v_i \in N(v_j) \end{cases}$$

Note that the matrix we just computed is symmetric. We easily see that the constant function mapping every vertex to 1, and represented by the coordinate vector $(1, 1, \cdots, 1)^T$, is in the kernel of $(T^* \circ T)$. Thereupon, it is an eigenfunction of the Laplacian with eigenvalue 0.

We claim that the relation between the multiplicity of 0 as an eigenvalue of the Laplacian and the connectivity of G goes as follows:

Proposition. *If G has k connected components, then 0 has multiplicity k as an eigenvalue of the Laplacian of G.*

Beweis. Denote $\Omega_1, \ldots, \Omega_k$ the connected components of the graph G. For $j \in \{1, \ldots, k\}$, let $\mathbf{1}_{\Omega_j}(x) \in \mathbb{R}^V$ denote the characteristic function of Ω_j. Since connected components are mutually disjoint, you can easily prove that these functions form a linearly independent subset of $\mathbb{R}^V$. We now show that every $\mathbf{1}_{\Omega_j}(x)$ is an eigenfunction of $T^* \circ T$ with eigenvalue 0. We have

$$(T^* \circ T)(\mathbf{1}_{\Omega_j})(x) = -2d\mathbf{1}_{\Omega_j}(x) - 2 \sum_{w \in N(x)} \mathbf{1}_{\Omega_j}(w).$$

Now, if $x \in \Omega_j$, this vanishes since neighbours share the same connected component. On the other hand, if $x \notin \Omega_j$, both terms in the RHS above vanish for the same reason. Hence $(T^* \circ T)(\mathbf{1}_{\Omega_j})(x) = 0$ for all $x \in V$. This shows that 0 has multiplicativity at least k.

Now assume for a contradiction that $f \in \mathbb{R}^V$ is not spanned by $\{\Omega_j \mid j \in \{1, \ldots, k\}\}$, i.e. assume that f isn't constant on any connected component of G, but that it is an eigenfunction of $T^* \circ T$ with eigenvalue 0. Then

$$0 = \langle (T^* \circ T)(f), f \rangle_V = \langle Tf, Tf \rangle_E = \sum_{e \in E} (Tf(e))^2 = \sum_{e \in E} (f(e^{(2)}) - f(e^{(1)}))^2.$$

This implies $f(e^{(2)}) - f(e^{(1)}) = 0$, $\forall e \in E$, and therefore that f is constant on the endpoints of any edge of G. This in turn implies that f is constant on connected components of G, which is a contradiction. □

Single Choice. In each exercise, exactly one answer is correct.

1. For which $x \in \mathbb{C}$ is the matrix $A := \begin{pmatrix} x & -x \\ x & x \end{pmatrix}$ unitary?

   - [(a)] For all $x \in \mathbb{C}$ with $|x|^2 = \frac{1}{2}$.
   - (b) Exactly for $x = \frac{1}{\sqrt{2}}$.
   - (c) For all $x \in \mathbb{C}$ with $x = -\overline{x}$.
   - (d) For $x = 0$.

   Erklärung: We compute $AA^* = \left(\begin{smallmatrix} 2x\overline{x} & 0 \\ 0 & 2x\overline{x} \end{smallmatrix}\right) = 2 \cdot |x|^2 \cdot I_2$; so (a) holds.

2. Which set is a subspace of the $\mathbb{C}$-vector space $M_{n\times n}(\mathbb{C})$?

   - (a) The set of unitary $n \times n$ matrices.
   - (b) The set of self-adjoint $n \times n$ matrices.
   - [(c)] The set of symmetric $n \times n$ matrices.
   - (d) The set of normal $n \times n$ matrices.

   Erklärung: For all symmetric matrices A and B and complex numbers λ, it holds that $(\lambda A + B)^T = \lambda A^T + B^T = \lambda A + B$; hence, (c) is correct. However, the 1×1 matrix $A := (1)$ is unitary and self-adjoint, but its multiple $2iA = (2i)$ is neither; therefore, (a) and (b) are false. Finally, in general, the sum of two normal matrices is not normal.

Multiple Choice Fragen

1. Let A be a Hermitian matrix. Which statements are correct?

   $\boxed{\text{(a)}}$ $\mathrm{Tr}(A) \in \mathbb{R}$.

   $\boxed{\text{(b)}}$ $\det(A) \in \mathbb{R}$.

   Explanation:

   (a) We have $a_{ii} = \overline{a_{ii}} \in \mathbb{R}$, therefore $\mathrm{tr}A = \sum a_{ii} \in \mathbb{R}$.

   (b) $\det A = \det\left(\overline{A^T}\right) = \overline{\det(A^T)} = \overline{\det A} \in \mathbb{R}$. Aliter: The eigenvalues λ_i of A are real. Therefore $\det(A) = \prod \lambda_i$ is also real. Hence $\mathrm{tr}A = \sum \lambda_i$.