Exercise 1 (Unit Balls of the Manhattan Norm) **2+1+2=5 points**

The *Manhattan norm* (or ℓ_1-norm) of a vector $\mathbf{x} = (x_1, \dots, x_\ell)^\intercal \in \mathbb{R}^\ell$ is defined as

$$\|\mathbf{x}\|_1 \coloneqq \sum_{i=1}^{\ell} |x_i| .$$

The ℓ-dimensional ℓ_1 *unit ball* is defined as $B_1^\ell \coloneqq \{\mathbf{x} \in \mathbb{R}^\ell \mid \|\mathbf{x}\|_1 \leq 1\}$.

a) (i) Draw $B_1^2 \subseteq \mathbb{R}^2$ in the plane.

(ii) Describe the shape of $B_1^3 \subseteq \mathbb{R}^3$.

b) Compute $\mathrm{vol}\left(B_1^2\right)$ and $\mathrm{vol}\left(B_1^3\right)$.

c) Prove that $\lim_{\ell\to\infty} \mathrm{vol}\left(B_1^\ell\right) = 0$.

Solution:

a) (i) B_1^2 is a (rotated) square with corner points $\binom{0}{1}$, $\binom{1}{0}$, $\binom{-1}{0}$, $\binom{0}{-1}$.

(ii) B_2^3 is an octahedron (double pyramid).

b) The volume of B_1^2 is the area of a square with side length $\sqrt{2}$ and thus, $\mathrm{vol}(B_1^2) = 2$.

The volume of B_1^3 is twice the volume of a pyramid whose base is a square of side length $\sqrt{2}$ and whose height is 1. The volume of a pyramid is given by $\frac{1}{3} \cdot B \cdot h$ where B is the base area and h the height. Thus, $\mathrm{vol}(B_1^3) = 2 \cdot \frac{1}{3} \cdot 2 \cdot 1 = \frac{4}{3}$.

(The general formula is $\mathrm{vol}(B_1^\ell) = \frac{2^\ell}{\ell!}$.)

c) For all $\mathbf{x} \in \mathbb{R}^\ell$ it holds that

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{\ell} |x_i| = \sqrt{\left(\sum_{i=1}^{\ell} |x_i|\right)^2} \geq \sqrt{\sum_{i=1}^{\ell} |x_i|^2} = \sqrt{\sum_{i=1}^{\ell} x_i^2} = \|\mathbf{x}\|.$$

Thus $B_1^\ell \subseteq B_2^\ell$ for all $\ell \geq 1$ where B_2^ℓ denotes the ℓ-dimensional ℓ_2 unit ball. Hence, $\mathrm{vol}(B_1^\ell) \leq \mathrm{vol}(B_2^\ell)$.

Since we know from the lecture that $\lim_{\ell\to\infty} \mathrm{vol}(B_2^\ell) = 0$, the claim follows.

Exercise 2 (Hyperball and Hypercube) **2+5=7 points**

Let $\ell \in \mathbb{N}_{>0}$ and let $s \in \mathbb{R}_{>0}$. We define $Q_{\ell,s}$ as the ℓ-dimensional *hypercube* of side length $s \in \mathbb{R}$ that is centred in the origin. That is,

$$Q_{\ell,s} = \left\{ (x_1, \dots, x_\ell)^\intercal \in \mathbb{R}^\ell \;\middle|\; |x_i| \leq \tfrac{s}{2} \text{ for all } i = 1, \dots, \ell \right\} = \left[-\tfrac{s}{2}, \tfrac{s}{2} \right]^\ell.$$

Note that $Q_{\ell,s}$ has 2^ℓ corners. We fill $Q_{\ell,s}$ with (Euclidean) hyperballs the following way:

1) We place 2^ℓ hyperballs of radius $\frac{s}{4}$ as close as possible to the 2^ℓ corners of the hypercube, so that their distance to the origin (i. e. the centre of $Q_{\ell,s}$) is maximal while still being completely contained in Q.

2) We place an additional single hyperball $B(Q_{\ell,s})$ in the origin (i. e. the centre of $Q_{\ell,s}$), such that its radius is maximal with the property that it intersects with none of the other hyperballs' interiors.

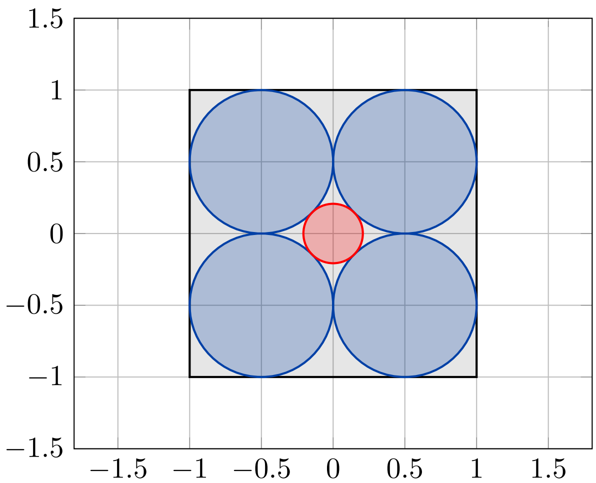
Solve the following tasks.

a) Sketch the situation (that is, the hypercube and all the hyperballs) for $\ell = 2$ and $s = 2$.

b) Let $s \in \mathbb{R}_{>0}$ be arbitrary but fixed. Find the minimal dimension $\ell \in \mathbb{N}$ for which it holds that $B(Q_{\ell,s}) \not\subseteq Q_{\ell,s}$ (that is, for which the final hyperball contains points outside of the hypercube). Prove that your answer is correct.

Solution:

a)



b) The 2^ℓ hyperballs from step 1 have centres

$$\left\{ \mathbf{c}_1, \dots, \mathbf{c}_{2^\ell} \right\} = \left\{ -\tfrac{s}{4}, \tfrac{s}{4} \right\}^\ell$$

and radius $\frac{s}{4}$. Let r be the radius of the final hyperball $B \coloneqq B(Q_{\ell,s})$ that is placed in the origin. Recall that r is chosen such that B does not intersect with the interiors of the other balls. This is equivalent to

$$r + \tfrac{s}{4} \leq \|\mathbf{c}\| = \sqrt{c_1^2 + \cdots + c_\ell^2} = \sqrt{\ell \cdot \left(\tfrac{s}{4}\right)^2} = \sqrt{\ell} \cdot \tfrac{s}{4}$$

for all $\mathbf{c} \in \{\mathbf{c}_1, \ldots, \mathbf{c}_{2^\ell}\}$. Since r was chosen maximal, it follows that $r = \frac{s}{4}(\sqrt{\ell} - 1)$. It holds that $B \not\subseteq Q_{\ell,s}$ if and only if $r > \frac{s}{2}$. Thus,

$$\tfrac{s}{2} < r = \tfrac{s}{4}(\sqrt{\ell} - 1) \quad \Leftrightarrow \quad 2 < \sqrt{\ell} - 1 \quad \Leftrightarrow \quad \ell > 9.$$

Thus, $\ell = 10$ is the minimal dimension for B to peek through the boundary of $Q_{\ell,s}$.

Exercise 3 (Power Iteration) **2+2+2+2=8 points**

Power iteration works since the successive multiplication with the same matrix shifts a randomly generated vector towards the eigenvector which belongs to the eigenvalue of the largest magnitude (the dominant eigenvalue). The algorithm starts with a random vector and terminates when this vector does not change anymore.
For the computations of the power iteration, we always start with the appropriate vector $\mathbf{x}$ consisting only of ones. For three dimensions this is $\mathbf{x} = (1, 1, 1)^T$.

Hint: You do not need to hand in any code, if used. It suffices to give the results up to 3 significant digits, for task a)-c), and 4 significant digits for task d).

a) Compute the power iteration on M_1 and M_2 for 5 iterations.

b) Compute three iterations of the power iteration procedure for M_3. Will the Power Iteration converge? If not, why does Power Iteration fail on this matrix? Justify your answer.

c) Observe that, if A is non-singular, then from $A\mathbf{x} = \lambda\mathbf{x}$ we get $A^{-1}\mathbf{x} = \frac{1}{\lambda}\mathbf{x}$. Use this to compute the eigenvalue with the smallest magnitude and corresponding eigenvector of M_1. What is the exact result?

d) How many iterations does the algorithm need until the eigenvector becomes stable for up to 3 significant digits for the matrices M_4 and M_5? Make sure to give your results with a precision of 4 significant digits and only stop when the first 3 digits (rounded correctly) become stable.

$$M_1 = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix};\ M_2 = \begin{pmatrix} 1 & 2 & 0 \\ -2 & 1 & 2 \\ 1 & 3 & 1 \end{pmatrix};\ M_3 = \begin{pmatrix} 1 & 1 & 0 \\ 3 & -1 & 0 \\ 0 & 0 & -2 \end{pmatrix};\ M_4 = \begin{pmatrix} 4 & 5 \\ 6 & 5 \end{pmatrix};\ M_5 = \begin{pmatrix} -4 & 10 \\ 7 & 5 \end{pmatrix}$$

Solution:

a) Power iteration for M_1:

$$\begin{pmatrix} -0.928 \\ -0.371 \end{pmatrix}, \begin{pmatrix} 0.942 \\ 0.336 \end{pmatrix}, \begin{pmatrix} -0.946 \\ -0.325 \end{pmatrix}, \begin{pmatrix} 0.947 \\ 0.320 \end{pmatrix}, \begin{pmatrix} -0.948 \\ -0.318 \end{pmatrix}$$

Power iteration for M_2:

$$\begin{pmatrix} 0.507 \\ 0.169 \\ 0.845 \end{pmatrix}, \begin{pmatrix} 0.382 \\ 0.382 \\ 0.841 \end{pmatrix}, \begin{pmatrix} 0.391 \\ 0.443 \\ 0.807 \end{pmatrix}, \begin{pmatrix} 0.411 \\ 0.411 \\ 0.814 \end{pmatrix}, \begin{pmatrix} 0.410 \\ 0.405 \\ 0.817 \end{pmatrix}$$

Note that M_1 has an eigenvalue -2 with eigenvector $(3, 1)^T$ and M_1 has an eigenvalue 3 with eigenvector $(1, 1, 2)^T$.

b) The three iterations give vectors $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ where $\mathbf{x}_0 = \frac{1}{\sqrt{3}}(1,1,1)^T$ and $\mathbf{x}_{i+1} = \frac{1}{\|M\mathbf{x}_i\|_2} M\mathbf{x}_i$. We have

$$\mathbf{x}_0 \approx \begin{pmatrix} 0.577 \\ 0.577 \\ 0.577 \end{pmatrix}, \mathbf{x}_1 \approx \begin{pmatrix} 0.577 \\ 0.577 \\ -0.577 \end{pmatrix}, \mathbf{x}_2 \approx \begin{pmatrix} 0.577 \\ 0.577 \\ 0.577 \end{pmatrix}, \mathbf{x}_3 \approx \begin{pmatrix} 0.577 \\ 0.577 \\ -0.577 \end{pmatrix}$$

In general, $\mathbf{x}_{2i} = \frac{1}{\sqrt{3}}(1,1,1)^T$ and $\mathbf{x}_{2i+1} = \frac{1}{\sqrt{3}}(1,1,-1)^T$ for all $i \geq 0$. Hence, the Power Iteration does not converge. The problem is that $(1,1,0)^T$ is an eigenvector with eigenvalue 2 and $(0,0,1)^T$ is an eigenvector with eigenvalue -2. In particular, $|\lambda_1| = |\lambda_2|$. The third eigenvector is $(-1,3,0)^T$ with eigenvalue -2.

c) We have

$$M_1^{-1} = \begin{pmatrix} -2.5 & 6 \\ -0.5 & 1 \end{pmatrix}$$

Performing the Power Iteration algorithm for 8 iterations we obtain that

$$\hat{\mathbf{x}}_{10} \approx \begin{pmatrix} 0.970 \\ 0.242 \end{pmatrix}.$$

This suggests that $(4,1)^T$ may be an eigenvector of M_1^{-1} which indeed it is, with eigenvalue -1. Hence, $(4,1)^T$ is also an eigenvector with eigenvalue -1 of the matrix M_1.

The results of all 8 rounds: [0.990 0.141], [-0.977 -0.212], [0.973 0.230], [-0.972 -0.237], [0.971 0.240], [-0.970 -0.241], [0.970 0.242], [-0.970 -0.242]

d) For M_4 it only takes 3 iterations until the eigenvector becomes stable for up to 3 significant digits. For M_5 it takes much longer, we require 61 iterations.

The results of 5 rounds for M_4:
1: [0.6332 0.7740] 2: [0.6409 0.7676] 3: [0.6401 0.7683] 4: [0.6402 0.7682] 5: [0.6402 0.7682]

The results of 65 rounds for M_5:
1: [0.4472 0.8944] 2: [0.6854 0.7282] 3: [0.4738 0.8806] 4: [0.6670 0.7451]
5: [0.4951 0.8689] 6: [0.6516 0.7586] 7: [0.5120 0.8590] 8: [0.6388 0.7694]
9: [0.5255 0.8508] 10: [0.6283 0.7780] 11: [0.5364 0.8440] 12: [0.6196 0.7849]
13: [0.5451 0.8384] 14: [0.6124 0.7905] 15: [0.5521 0.8338] 16: [0.6066 0.7950]
17: [0.5577 0.8301] 18: [0.6019 0.7986] 19: [0.5622 0.8270] 20: [0.5980 0.8015]
21: [0.5659 0.8245] 22: [0.5948 0.8038] 23: [0.5688 0.8225] 24: [0.5923 0.8057]
25: [0.5712 0.8208] 26: [0.5902 0.8073] 27: [0.5731 0.8195] 28: [0.5885 0.8085]
29: [0.5747 0.8184] 30: [0.5871 0.8095] 31: [0.5759 0.8175] 32: [0.5860 0.8103]
33: [0.5769 0.8168] 34: [0.5851 0.8110] 35: [0.5777 0.8162] 36: [0.5844 0.8115]
37: [0.5784 0.8157] 38: [0.5838 0.8119] 39: [0.5790 0.8154] 40: [0.5833 0.8123]
41: [0.5794 0.8151] 42: [0.5829 0.8125] 43: [0.5797 0.8148] 44: [0.5826 0.8128]
45: [0.5800 0.8146] 46: [0.5823 0.8130] 47: [0.5803 0.8144] 48: [0.5821 0.8131]

49: [0.5804 0.8143] 50: [0.5820 0.8132] 51: [0.5806 0.8142] 52: [0.5818 0.8133]
53: [0.5807 0.8141] 54: [0.5817 0.8134] 55: [0.5808 0.8140] 56: [0.5816 0.8135]
57: [0.5809 0.8140] 58: [0.5815 0.8135] 59: [0.5810 0.8139] 60: [0.5815 0.8136]
61: [0.5810 0.8139] 62: [0.5814 0.8136] 63: [0.5811 0.8139] 64: [0.5814 0.8136]
65: [0.5811 0.8138]

Exercise 4 (Positive Semi-Definite Matrices) 0 points

This exercise will not be corrected and awards 0 points.

A symmetric matrix $A \in \mathbb{R}^{n\times n}$ is called *positive semi-definite* if every eigenvalue of A is non-negative. Solve the following tasks.

a) For $c \in \mathbb{R}$, consider the matrices

$$A_c = \begin{pmatrix} 2 & 0 & c \\ 0 & 1 & 0 \\ c & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Determine the set $\{c \in \mathbb{R} \mid A_c \text{ is positive semi-definite}\}$. Give the answer as an interval in $\mathbb{R}$ and prove that it is correct.

b) Prove that the following is an equivalent definition to the one given above: A symmetric matrix $A \in \mathbb{R}^{n\times n}$ is positive semi-definite if and only if there exists $B \in \mathbb{R}^{n\times n}$ such that $A = BB^{\intercal}$.

c) Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix, let $\lambda_1 \neq \lambda_2$ be two distinct eigenvalues of A and let E_1 and E_2 denote the corresponding eigenspaces. Prove that E_1 and E_2 are orthogonal, that is, for all $\mathbf{v}_1 \in E_1$ and $\mathbf{v}_2 \in E_2$ it holds that $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$.

Solution:

a) The characteristic polynomial is given by

$$\begin{aligned}
\det\big(A_c - \lambda I\big) &= \det \begin{pmatrix} 2-\lambda & 0 & c \\ 0 & 1-\lambda & 0 \\ c & 0 & 1-\lambda \end{pmatrix} \\
&= (2-\lambda)\cdot\big((1-\lambda)^2 - 0\big) - 0 + c\big(0 - c(1-\lambda)\big) \\
&= (1-\lambda)\big((2-\lambda)(1-\lambda) - c^2\big) \\
&= (1-\lambda)\big(\lambda^2 - 3\lambda + (2-c^2)\big).
\end{aligned}$$

Thus, for all $c \in \mathbb{R}$, $\lambda_1 = 1$ is an eigenvalue of A_c. The other two eigenvalues are

$$\lambda_{2,3} = \tfrac{3}{2} \pm \sqrt{\left(\tfrac{3}{2}\right)^2 - 2 + c^2} = \frac{3 \pm \sqrt{4c^2+1}}{2}.$$

One of these, say λ_2, is at least $\frac{3}{2}$, so in particular greater than 0. For the other one, λ_3, it holds that

$$\lambda_3 \geq 0 \quad \Leftrightarrow \quad \sqrt{4c^2+1} \leq 3 \quad \Leftrightarrow \quad |c| \leq \sqrt{2}.$$

Thus, A_c is positive semi-definite if and only if $c \in [-\sqrt{2}, \sqrt{2}]$.

b) Let $A \in \mathbb{R}^{n\times n}$ be symmetric. Note that all eigenvalues of A are real.

First suppose $A = BB^\intercal$ for some $B \in \mathbb{R}^{n\times n}$ and let $(\lambda, \mathbf{v})$ be an eigenpair of A. Then

$$\lambda\langle \mathbf{v}, \mathbf{v}\rangle = \mathbf{v}^\intercal A\mathbf{v} = \mathbf{v}^\intercal BB^\intercal \mathbf{v} = \langle B\mathbf{v}, B\mathbf{v}\rangle.$$

Since $\langle \mathbf{v}, \mathbf{v}\rangle \geq 0$ and $\langle B\mathbf{v}, B\mathbf{v}\rangle \geq 0$ it follows that $\lambda \geq 0$.

For the other direction suppose that all eigenvalues of A are non-negative. Since A is symmetric, by the Spectral Decomposition Theorem, it can be written as $A = U\Lambda U^\intercal$ for some orthogonal $U \in \mathbb{R}^{n\times n}$ and $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ where $\lambda_1, \ldots, \lambda_n$ are the (non-negative) eigenvalues of A. Define $S = \operatorname{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$. Then $S = S^\intercal$ and $\Lambda = SS = SS^\intercal$. So

$$A = U(SS^\intercal)U^\intercal = (US)(US)^\intercal.$$

c) Let $\mathbf{v}_1 \in E_1$ and $\mathbf{v}_2 \in E_2$ be arbitrary but fixed. Then

$$\mathbf{v}_1{}^\intercal A\mathbf{v}_2 = \mathbf{v}_1{}^\intercal(A\mathbf{v}_2) = \mathbf{v}_1{}^\intercal \lambda_2\mathbf{v}_2 = \lambda_2\langle \mathbf{v}_1, \mathbf{v}_2\rangle$$

and

$$\mathbf{v}_1{}^\intercal A\mathbf{v}_2 = (A^\intercal \mathbf{v}_1)^\intercal \mathbf{v}_2 = (A\mathbf{v}_1)^\intercal \mathbf{v}_2 = (\lambda_1\mathbf{v}_1)^\intercal \mathbf{v}_2 = \lambda_1\langle \mathbf{v}_1, \mathbf{v}_2\rangle.$$

That is, $\lambda_1\langle \mathbf{v}_1, \mathbf{v}_2\rangle = \lambda_2\langle \mathbf{v}_1, \mathbf{v}_2\rangle$. As $\lambda_1 \neq \lambda_2$, this implies $\langle \mathbf{v}_1, \mathbf{v}_2\rangle = 0$.