2. Vectors

Exercise 2.1 (Subpaces and dimensions) Consider the set $\mathcal{S}$ of points such that

$$x_1 + 2x_2 + 3x_3 = 0, \;\; 3x_1 + 2x_2 + x_3 = 0.$$

Show that $\mathcal{S}$ is a subspace. Determine its dimension, and find a basis for it.

Solution 2.1 The set $\mathcal{S}$ is a subspace, as can be checked directly: if $x, y \in \mathcal{S}$, then for every $\lambda, \mu \in \mathbb{R}$, we have $\lambda x + \mu y \in \mathcal{S}$. To find the dimension, we solve the equation and find that any solution to the equations is of the form $x_1 = -1/2x_2$, $x_3 = -1/3x_2$, where x_2 is free. Hence the dimension of $\mathcal{S}$ is 1, and a basis for S is the vector $(-1/2, 1, -1/3)$.

Exercise 2.2 (Affine sets and projections) Consider the set in $\mathbb{R}^3$, defined by the equation

$$\mathcal{P} = \left\{ x \in \mathbb{R}^3 \; : \; x_1 + 2x_2 + 3x_3 = 1 \right\}.$$

1. Show that the set $\mathcal{P}$ is an affine set of dimension 2. To this end, express it as $x^{(0)} + \operatorname{span}(x^{(1)}, x^{(2)})$, where $x^{(0)} \in \mathcal{P}$, and $x^{(1)}, x^{(2)}$ are linearly independent vectors.

2. Find the minimum Euclidean distance from 0 to the set $\mathcal{P}$, and a point that achieves the minimum distance.

Solution 2.2

1. We can express any vector $x \in \mathcal{P}$ as $x = (x_1, x_2, 1/3 - x_1/3 - 2x_2/3)$, where x_1, x_2 are arbitrary. Thus

   $$x = x^{(0)} + x_1 x^{(1)} + x_2 x^{(2)},$$

   where

   $$x^{(0)} = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{3} \end{bmatrix}, \;\; x^{(1)} = \begin{bmatrix} 1 \\ 0 \\ -\frac{1}{3} \end{bmatrix}, \;\; x^{(2)} = \begin{bmatrix} 0 \\ 1 \\ -\frac{2}{3} \end{bmatrix}.$$

   Since $x^{(1)}$ and $x^{(2)}$ are linearly independent, $\mathcal{P}$ is of dimension 2.

2. The set $\mathcal{P}$ is defined by a single linear equation $a^\top x = b$, with $a^\top = [1 \; 2 \; 3]$ and $b = 1$, i.e., $\mathcal{P}$ is a hyperplane. The minimum Euclidean distance from o to $\mathcal{P}$ is the ℓ_2 norm of the projection of 0 onto $\mathcal{P}$, which can be determined as discussed in Section 2.3.2.2. That is, the projection x^* of 0 onto $\mathcal{P}$ is such that $x^* \in \mathcal{P}$ and x^*

is orthogonal to the subspace generating $\mathcal{P}$ (which coincides with the span of a), that is $x^* = \alpha a$. Hence, it must be that $a^\top x^* = 1$, thus $\alpha\|a\|_2^2 = 1$, and $\alpha = 1/\|a\|_2^2$. We thus have that

$$x^* = \frac{a}{\|a\|_2^2},$$

and the distance we are seeking is $\|x^*\|_2 = 1/\|a\|_2 = 1/\sqrt{14}$.

Exercise 2.3 (Angles, lines and projections)

1. Find the projection z of the vector $x = (2,1)$ on the line that passes through $x_0 = (1,2)$ and with direction given by vector $u = (1,1)$.

2. Determine the angle between the following two vectors:

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad y = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}.$$

Are these vectors linearly independent?

Solution 2.3

1. We can observe directly that $u^\top(x - x_0) = 0$, hence the projection of x is the same as that of x_0, which is $z = x_0$ itself.

Alternatively, as seen in Section (2.3.2.1), the projection is

$$z = x_0 + \frac{u^\top(x - x_0)}{u^\top u} u$$

which gives $z = x_0$.

Another method consists in solving

$$\begin{aligned} \min_t \|x_0 + tu - x\|_2^2 &= \min_t t^2 u^T u - 2tu^\top(x - x_0) + \|x - x_0\|_2^2 \\ &= \min_t (u^\top u)(t - t_0)^2 + \text{constant}, \end{aligned}$$

where $t_0 = (x - x_0)^\top u/(u^T u)$. This leads to the optimal $t^* = t_0$, and provides the same result as before.

2. The angle cosine is given by

$$\cos\theta = \frac{x^\top y}{\|x\|_2 \|y\|_2} = \frac{10}{14},$$

which gives $\theta \approx 41°$.

The vectors are linearly independent, since $\lambda x + \mu y = 0$ for $\lambda, \mu \in \mathbb{R}$ implies that $\lambda = \mu = 0$. Another way to prove this is to observe that the angle is not $0°$ nor $180°$.

Exercise 2.4 (Inner product) Let $x, y \in \mathbb{R}^n$. Under which condition on $\alpha \in \mathbb{R}^n$ does the function

$$f(x,y) = \sum_{k=1}^{n} \alpha_k x_k y_k$$

define an inner product on $\mathbb{R}^n$?

Solution 2.4 The axioms of 2.2 are all satisfied for any $\alpha \in \mathbb{R}^n$, except the conditions

$$\begin{aligned} &f(x,x) \geq 0; \\ &f(x,x) = 0 \text{ if and only if } x = 0. \end{aligned}$$

These properties hold if and only if $\alpha_k > 0$, $k = 1, \ldots, n$. Indeed, if the latter is true, then the above two conditions hold. Conversely, if if there exist k such that $\alpha_k \leq 0$, setting $x = e_k$ (the k-th unit vector in $\mathbb{R}^n$) produces $f(e_k, e_k) \leq 0$; this contradicts one of the two above conditions.

Exercise 2.5 (Orthogonality) Let $x, y \in \mathbb{R}^n$ be two unit-norm vectors, that is, such that $\|x\|_2 = \|y\|_2 = 1$. Show that the vectors $x - y$ and $x + y$ are orthogonal. Use this to find an orthogonal basis for the subspace spanned by x and y.

Solution 2.5 When x, y are both unit-norm, we have

$$(x-y)^\top (x+y) = x^\top x - y^\top y - y^\top x + x^\top y = x^\top x - y^\top y = 0,$$

as claimed.

We can express any vector $z \in \text{span}(x, y)$ as $z = \lambda x + \mu y$, for some $\lambda, \mu \in \mathbb{R}$. We have $z = \alpha u + \beta v$, where

$$\alpha = \frac{\lambda + \mu}{2}, \quad \beta = \frac{\lambda - \mu}{2}.$$

Hence $z \in \text{span}(u, v)$. The converse is also true for similar reasons. Thus, (u, v) is an orthogonal basis for $\text{span}(x, y)$. We finish by normalizing u, v, replacing them with $(u/\|u\|_2, v/\|v\|_2)$. The desired orthogonal basis is thus given by $((x-y)/\|x-y\|_2, (x+y)/\|x+y\|_2)$.

Exercise 2.6 (Norm inequalities)

1. Show that the following inequalities hold for any vector x:

$$\frac{1}{\sqrt{n}}\|x\|_2 \leq \|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2 \leq n\|x\|_\infty.$$

Hint: use the Cauchy-Schwartz inequality.

2. Show that for any non-zero vector x,

$$\mathrm{card}(x) \geq \frac{\|x\|_1^2}{\|x\|_2^2},$$

where $\mathrm{card}(x)$ is the *cardinality* of the vector x, defined as the number of non-zero elements in x. Find vectors x for which the lower bound is attained.

Solution 2.6

1. We have

$$\|x\|_2^2 = \sum_{i=1}^n x_i^2 \leq n \cdot \max_i x_i^2 = n \cdot \|x\|_\infty^2.$$

Also, $\|x\|_\infty \leq \sqrt{x_1^2 + \ldots + x_n^2} = \|x\|_2$.

The inequality $\|x\|_2 \leq \|x\|_1$ is obtained after squaring both sides, and checking that

$$\sum_{i=1}^n x_i^2 \leq \sum_{i=1}^n x_i^2 + \sum_{i \neq j} |x_i x_j| = \left(\sum_{i=1}^n |x_i|\right)^2 = \|x\|_1^2.$$

Finally, the condition $\|x\|_1 \leq \sqrt{n}\|x\|_2$ is due to the Cauchy-Schwartz inequality

$$|z^\top y| \leq \|y\|_2 \cdot \|z\|_2,$$

applied to the two vectors $y = (1, \ldots, 1)$ and $z = |x| = (|x_1|, \ldots, |x_n|)$.

2. Let us apply the Cauchy-Schwartz inequality with $z = |x|$ again, and with y a vector with $y_i = 1$ if $x_i \neq 0$, and $y_i = 0$ otherwise. We have $\|y\|_2 = \sqrt{k}$, with $k = \mathrm{card}(x)$. Hence

$$|z^\top y| = \|x\|_1 \leq \|y\|_2 \cdot \|z\|_2 = \sqrt{k} \cdot \|x\|_2,$$

which proves the result. The bound is attained for vectors with k non-zero elements, all with the same magnitude.

Exercise 2.7 (Hölder inequality) Prove Hölder's inequality (2.4). *Hint:* consider the normalized vectors $u = x/\|x\|_p$, $v = y/\|y\|_q$, and observe that

$$|x^\top y| = \|x\|_p\|y\|_q \cdot |u^\top v| \leq \|x\|_p\|y\|_q \sum_k |u_k v_k|.$$

Then, apply Young's inequality (see Example 8.10) to the products $|u_k v_k| = |u_k||v_k|$.

Solution 2.7 The inequality is trivial if one of the vectors x, y is zero. We henceforth assume that none is, which allows us to define the normalized vectors u, v. We need to show that

$$\sum_k |u_k v_k| \leq 1.$$

Using the hint given, we apply Young's inequality, which states that for any given numbers $a, b \geq 0$ and $p, q > 0$ such that

$$\frac{1}{p} + \frac{1}{q} = 1,$$

it holds that

$$ab \leq \frac{1}{p} a^p + \frac{1}{q} b^q.$$

We thus have, with $a = |u_k|$ and $b = |v_k|$, and summing over k:

$$\begin{aligned} \sum_k |u_k v_k| &\leq \frac{1}{p} \sum_k |u_k|^p + \frac{1}{q} \sum_k |v_k|^q \\ &= \frac{1}{p} \|u\|_p^p + \frac{1}{q} \|v\|_q^q \\ &= \frac{1}{p} + \frac{1}{q} = 1, \end{aligned}$$

where we have used the fact that $\|u\|_p = \|v\|_q = 1$.

Exercise 2.8 (Linear functions)

1. For a n-vector x, with $n = 2m - 1$ odd, we define the median of x as the scalar value x_a such that exactly n of the values in x are $\leq x_a$ and n are $\geq x_a$ (i.e., x_a leaves half of the values in x to its left, and half to its right). Now consider the function $f : \mathbb{R}^n \to \mathbb{R}$, with values $f(x) = x_a - \frac{1}{n} \sum_{i=1}^n x_i$. Express f as a scalar product, that is, find $a \in \mathbb{R}^n$ such that $f(x) = a^\top x$ for every x. Find a basis for the set of points x such that $f(x) = 0$.

2. For $\alpha \in \mathbb{R}^2$, we consider the "power law" function $f : \mathbb{R}^2_{++} \to \mathbb{R}$, with values $f(x) = x_1^{\alpha_1} x_2^{\alpha_2}$. Justify the statement: "the coefficients α_i provide the ratio between the relative error in f to a relative error in x_i".

Solution 2.8 (Linear functions) TBD

Exercise 2.9 (Bound on a polynomial's derivative) In this exercise, you derive a bound on the largest absolute value of the derivative of a polynomial of a given order, in terms of the size of the coefficients[1]. For $w \in \mathbb{R}^{k+1}$, we define the polynomial p_w, with values

[1] See the discussion on regularization in Section 13.2.3 for an application of this result.

$$p_w(x) \doteq w_1 + w_2 x + \ldots + w_{k+1} x^k.$$

Show that, for any $p \geq 1$

$$\forall\, x \in [-1, 1] \; : \; \left| \frac{\mathrm{d}p_w(x)}{\mathrm{d}x} \right| \leq C(k, p) \|v\|_p,$$

where $v = (w_2, \ldots, w_{k+1}) \in \mathbb{R}^k$, and

$$C(k, p) = \begin{cases} k & p = 1, \\ k^{3/2} & p = 2, \\ \frac{k(k+1)}{2} & p = \infty. \end{cases}$$

Hint: you may use Hölder's inequality (2.4) or the results from Exercise 2.6.

Solution 2.9 (Bound on a polynomial's derivative) We have, with $z = (1, 2, \ldots, k)$, and using Hölder's inequality:

$$\begin{aligned} \left| \frac{\mathrm{d}p_w(x)}{\mathrm{d}x} \right| &= \left| w_2 + 2w_3 x + \ldots + k w_{k+1} x^{k-1} \right| \\ &\leq |w_2| + 2|w_3| + \ldots + k|w_{k+1}| \\ &= |v^\top z| \\ &\leq \|v\|_p \cdot \|z\|_q. \end{aligned}$$

When $p = 1$, we have

$$\|z\|_q = \|z\|_\infty = k.$$

When $p = 2$, we have

$$\|z\|_q = \|z\|_2 = \sqrt{1 + 4 + \ldots + k^2} \leq \sqrt{k \cdot k^2} = k^{3/2}.$$

When $p = \infty$, we have

$$\|z\|_q = \|z\|_1 = 1 + 2 + \ldots + k = \frac{k(k+1)}{2}.$$