Suggestions for midterm review #3

The repetitoria are usually not complete; I am merely bringing up the points that many people didn't know on the recitations.

0.1. Linear transformations

The following mostly comes from my recitation #11 (with minor changes) and so is unsuited for the exam. I would begin with repeating stuff, as people seem to be highly confused:

0.1.1. Repetitorium

1. A map $T : V \to W$ between two vector spaces (say, $\mathbb{R}$-vector spaces) is *linear* if and only if it satisfies the axioms

$$\begin{aligned} T(0) &= 0; \\ T(u+v) &= T(u) + T(v) \qquad \text{for all } u, v \in V; \\ T(\alpha u) &= \alpha T(u) \qquad \text{for all } u \in V \text{ and } \alpha \in \mathbb{R} \end{aligned}$$

(where the $\mathbb{R}$ should be a $\mathbb{C}$ if the vector spaces are complex). You can leave out the first axiom (it follows from applying the second axiom to $u = 0$ and $v = 0$), but the other two axioms are essential. Examples:

- The map $T : \mathbb{R}^2 \to \mathbb{R}^3$ sending every $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ to $\begin{pmatrix} x_1 \\ x_1 - 2x_2 \\ 3x_2 \end{pmatrix}$ is linear. (Indeed, for every $u \in \mathbb{R}^2$ and $v \in \mathbb{R}^2$, we can write $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ and $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, and then we have $T(u+v) = T\left(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right) = T\left(\begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}\right) = \begin{pmatrix} u_1 + v_1 \\ (u_1 + v_1) - 2(u_2 + v_2) \\ 3(u_2 + v_2) \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_1 + v_1 - 2u_2 - 2v_2 \\ 3u_2 + 3v_2 \end{pmatrix}$, which is the same as $T(u) + T(v) = T\left(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right) + T\left(\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right) = \begin{pmatrix} u_1 \\ u_1 - 2u_2 \\ 3u_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_1 - 2v_2 \\ 3v_2 \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_1 + v_1 - 2u_2 - 2v_2 \\ 3u_2 + 3v_2 \end{pmatrix}$. This proves the second axiom. The other axioms are just as easy to check.)

- The map $T : \mathbb{R}^2 \to \mathbb{R}$ sending every $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ to $x_1 + x_2 - 1$ is not linear. (Indeed, it fails the $T(0) = 0$ axiom. It also fails the other two axioms, but failing one of them is enough for it to be not linear.)

- The map $T : \mathbb{R} \to \mathbb{R}^2$ sending every x to $\begin{pmatrix} x \\ x^2 \end{pmatrix}$ is not linear. (Indeed, it fails the second axiom for $u = 1$ and $v = 1$ because $(1+1)^2 \neq 1^2 + 1^2$.)

2. If V and W are two vector spaces, and if $T : V \to W$ is a linear map, then the *matrix representation* of T with respect to a given basis $(v_1, v_2, \ldots, v_n)$ of V and a given basis $(w_1, w_2, \ldots, w_m)$ of W is the $m \times n$-matrix M_T defined as follows: For every $j \in \{1, 2, \ldots, n\}$, expand the vector $T(v_j)$ with respect to the basis $(w_1, w_2, \ldots, w_m)$, say, as follows:

$$T(v_j) = \alpha_{1,j} w_1 + \alpha_{2,j} w_2 + \cdots + \alpha_{m,j} w_m.$$

Then, M_T is the $m \times n$-matrix whose (i, j)-th entry is $\alpha_{i,j}$.

For example, if $n = 2$ and $m = 3$, then

$$M_T = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \\ \alpha_{3,1} & \alpha_{3,2} \end{pmatrix},$$

where

$$T(v_1) = \alpha_{1,1} w_1 + \alpha_{2,1} w_2 + \alpha_{3,1} w_3;$$
$$T(v_2) = \alpha_{1,2} w_1 + \alpha_{2,2} w_2 + \alpha_{3,2} w_3.$$

3. What is this matrix M_T good for? First of all, it allows easily expanding $T(v)$ in the basis $(w_1, w_2, \ldots, w_m)$ of W if v is a vector in V whose expansion in the basis $(v_1, v_2, \ldots, v_n)$ of V is known.[1] More importantly, once you know the matrix M_T, you can use it as a proxy whenever you want to know something about T. For instance, you want to know the nullspace of T. You know how to compute the nullspace of a matrix, but not how to compute the nullspace of a linear map. So you find the nullspace of M_T; it consists of vectors of the form $\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{pmatrix}$. Then, the nullspace of T consists of the corresponding vectors $\gamma_1 v_1 + \gamma_2 v_2 + \cdots + \gamma_n v_n$.

4. Suppose that $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$. These vector spaces already have standard bases, but nothing keeps you from representing a linear map $T : \mathbb{R}^n \to$

[1] For instance, if v is one of the basis vectors v_j, then the expansion of $T(v_j)$ can be simply read off from the j-th column of M_T; otherwise, it is an appropriate combination of columns.

$\mathbb{R}^m$ with respect to two **other** bases. There is a formula to represent the map in this case:

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map, and let A be the $m \times n$-matrix representation of T with respect to the standard bases (that is, the matrix such that $T(v) = Av$ for every $v \in \mathbb{R}^n$). Let $(v_1, v_2, \ldots, v_n)$ be a basis of $\mathbb{R}^n$, and let $(w_1, w_2, \ldots, w_m)$ be a basis of $\mathbb{R}^m$. Let B be the matrix with columns $v_1, v_2, \ldots, v_n$, and let C be the matrix with columns $(w_1, w_2, \ldots, w_m)$. Then, the matrix representation of T with respect to the bases $(v_1, v_2, \ldots, v_n)$ and $(w_1, w_2, \ldots, w_m)$ is $M_T = C^{-1}AB$.

0.1.2. Exercises

Exercise 0.1. Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear map which sends every vector $v \in \mathbb{R}^3$ to $\begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 1 \end{pmatrix} v$.

Consider the following basis (v_1, v_2) of the vector space $\mathbb{R}^2$:

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Consider the following basis (w_1, w_2, w_3) of the vector space $\mathbb{R}^3$:

$$w_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad w_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \qquad w_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Find the matrix M_f representing f with respect to these two bases (v_1, v_2) and (w_1, w_2, w_3).

[**Hint:** You can use the $M_T = C^{-1}AB$ formula here, with $B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ and $A = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 1 \end{pmatrix}$. But it is easier to do it directly, and the $C^{-1}AB$ formula does not help in the next exercises.]

First solution. We compute the matrix representation directly, using part **2.** of the repetitorium above. We have

$$f(v_1) = f\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix} = 0w_1 + 2w_2 + 2w_3$$

[2] and similarly

$$f\left(v_2\right) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = 1w_1 + 1w_2 + 1w_3.$$

Thus,

$$M_f = \begin{pmatrix} 0 & 1 \\ 2 & 1 \\ 2 & 1 \end{pmatrix}.$$

Second solution. We can apply part **4.** of the repetitorium. Here, $n = 2$, $m = 3$, $T = f$, $A = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, and $C = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$. Thus, $M_f = C^{-1}AB = \begin{pmatrix} 0 & 1 \\ 2 & 1 \\ 2 & 1 \end{pmatrix}$.

For every $n \in \mathbb{N}$, we let P_n denote the vector space of all polynomials (with real coefficients) of degree $\leq n$ in one variable x. This vector space has dimension $n+1$, and its simplest basis is $\left(1, x, x^2, \ldots, x^n\right)$. We call this basis the *monomial basis* of P_n.

Exercise 0.2. Which of the following maps are linear? For every one that is, represent it as a matrix with respect to the monomial bases of its domain and its target.

(a) The map $T_a : P_2 \to P_2$ given by $T_a(f) = f(x+1)$.
(b) The map $T_b : P_2 \to P_3$ given by $T_b(f) = xf(x)$.
(c) The map $T_c : P_2 \to P_4$ given by $T_c(f) = f(1)f(x)$.
(d) The map $T_d : P_2 \to P_4$ given by $T_d(f) = f\left(x^2+1\right)$.
(e) The map $T_e : P_2 \to P_2$ given by $T_e(f) = x^2 f\left(\dfrac{1}{x}\right)$.

---

[2]Finding the expression $0w_1 + 2w_2 + 0w_3$ for $\begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix}$ can be done by solving a system of linear equations (in general, if we want to represent a vector $p \in \mathbb{R}^n$ as a linear combination $p = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_n w_n$ of n given linearly independent vectors $(w_1, w_2, \ldots, w_n)$, then we can find the coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$ by solving the equation $W \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = p$, where W denotes the $n \times n$-matrix with columns $w_1, w_2, \ldots, w_n$). But in our case, the vectors w_1, w_2, w_3 are so simple that we can easily represent an arbitrary vector $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ as their linear combination, as follows: $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 w_1 + (x_2 - x_1) w_2 + (x_3 - x_2) w_3$.

(g) The map $T_g : P_3 \to P_2$ given by $T_g(f) = f'(x)$.

[There is no part **(f)** because I want to avoid having a T_f map when f stands for a polynomial.]

(h) The map $T_h : P_3 \to P_2$ given by $T_h(f) = f(x) - f(x-1)$.

(i) The map $T_i : P_3 \to P_2$ given by $T_i(f) = f(x) - 2f(x-1) + f(x-2)$.

Solution. First, wet us introduce a new notation to clear up some ambiguities. In the exercise, we have been writing $f(x+1)$ for "f with $x+1$ substituted for x". This notation could, however, be mistaken for the product of f with $x+1$. (This confusion is particularly bad when f is, say, x; normal people would read $x(x+1)$ as "$x \cdot (x+1)$" and not as "x with $x+1$ substituted for x".) So we change this notation to $f\mid_{x+1}$. More generally, if f is a polynomial and y is a number or another polynomial, we let $f\mid_y$ be f with y substituted for x. For instance,

$$\begin{aligned}
\left(x^2+x+1\right)\mid_3 &= 3^2+3+1; \\
\left(x^2+x+1\right)\mid_{x+1} &= (x+1)^2+(x+1)+1; \\
\left(x^3-3x+2\right)\mid_{x+1} &= (x+1)^3-3(x+1)+2; \\
\left(x^3-3x+2\right)\mid_{x^2+1} &= \left(x^2+1\right)^3-3\left(x^2+1\right)+2; \\
7\mid_{x^2+1} &= 7
\end{aligned}$$

(notice that constant polynomials do not change under substitution, because they have nothing to substitute in them). Now, the map T_a of part **(a)** is given by $T_a(f) = f\mid_{x+1}$, whereas the map T_b of part **(b)** is given by $T_b(f) = x \cdot (f\mid_x) = x \cdot f$ (clearly, $f\mid_x = f$, since we are substituting x for itself), and the map T_c of part **(c)** is given by $T_c(f) = (f\mid_1)(f\mid_x) = (f\mid_1) f$, and so on.

(a) The map T_a is linear. To prove this, we need to check that it satisfies the three axioms (see part **1.** of the repetitorium); let us only check the second one, leaving the others to the reader.

For any $u, v \in P_2$ (this means that u and v are two polynomials of degree ≤ 2), we have

$$T_a(u+v) = (u+v)\mid_{x+1} = u\mid_{x+1} + v\mid_{x+1}$$

(because substituting $x+1$ for x in a sum of two polynomials gives the same result as substituting $x+1$ for x in each of the polynomials and then adding the results). Comparing this with $\underbrace{T_a(u)}_{=u\mid_{x+1}} + \underbrace{T_a(v)}_{=v\mid_{x+1}} = u\mid_{x+1} + v\mid_{x+1}$, we obtain $T_a(u+v) = T_a(u) + T_a(v)$. This proves that the second axiom is satisfied. The other axioms, as already said, work similarly.

We are not done yet, as we still have to find the matrix representation M_{T_a} of T_a with respect to the monomial bases of P_2 and P_2. These bases are $\left(1, x, x^2\right)$

and $\left(1, x, x^2\right)$ (these are the same basis twice). To find the representation, we use part **2.** of the repetitorium above. We have

$$\begin{aligned} T_a(1) &= 1 \mid_{x+1} = 1 \qquad \left(\begin{array}{c}\text{since 1 is a constant polynomial and} \\ \text{thus unchanged under substitution}\end{array}\right) \\ &= 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2; \\ T_a(x) &= x \mid_{x+1} = x + 1 = 1 \cdot 1 + 1 \cdot x + 0 \cdot x^2; \\ T_a\left(x^2\right) &= x^2 \mid_{x+1} = (x+1)^2 = 1 \cdot 1 + 2 \cdot x + 1 \cdot x^2. \end{aligned}$$

Thus, the matrix M_{T_a} is

$$M_{T_a} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

(b) The map T_b is linear. Here (and in the following), we shall not give the proof, as it always proceeds straightforwardly along the lines of part **(a)**. (In algebra, you will eventually encounter linear maps whose linearity is difficult to prove; but you will not see them soon.)

To find its matrix representation M_{T_b} with respect to the monomial bases $\left(1, x, x^2\right)$ and $\left(1, x, x^2, x^3\right)$ of P_2 and P_3, we again use part **2.** of the repetitorium. We have

$$\begin{aligned} T_b(1) &= x \underbrace{(1 \mid_x)}_{=1} = x1 = x = 0 \cdot 1 + 1 \cdot x + 0 \cdot x^2 + 0 \cdot x^3; \\ T_b(x) &= x \underbrace{(x \mid_x)}_{=x} = xx = x^2 = 0 \cdot 1 + 0 \cdot x + 1 \cdot x^2 + 0 \cdot x^3; \\ T_b\left(x^2\right) &= x \underbrace{\left(x^2 \mid_x\right)}_{=x^2} = xx^2 = x^3 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + 1 \cdot x^3. \end{aligned}$$

Thus, the matrix M_{T_b} is

$$M_{T_b} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(c) The map T_c is **not** linear, as witnessed, e.g., by the failure of its second axiom on $u = 2$ and $v = 2$ (indeed, $T_c(2+2) = \underbrace{((2+2) \mid_1)}_{=4|_1=4} \underbrace{((2+2) \mid_x)}_{=4|_x=4} = 4 \cdot 4 = 16$ differs from $T_c(2) + T_c(2) = 8$). Therefore the question of a matrix representation is moot.[3]

[3]You **could** follow the procedure of part **2.** of the repetitorium to obtain a matrix M_{T_c}, but this matrix will not be a "matrix representation" for T_c; it only knows about the values of T_c on the basis elements $1, x, x^2$, and (absent linearity of T_c) says nothing about how T_c behaves on the rest of the space P_2.

(d) The map T_d is linear (the proof is similar to that in part **(a)**). The matrix M_{T_d} (which can be obtained in the same way as M_{T_a} and M_{T_b}) is

$$M_{T_d} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(For example, its third column comes from the fact that

$$T_d\left(x^2\right) = x^2 \mid_{x^2+1} = \left(x^2+1\right)^2 = x^4 + 2x^2 + 1 = 1 \cdot 1 + 0 \cdot x + 2 \cdot x^2 + 0 \cdot x^3 + 1 \cdot x^4.$$

)

(e) The map T_e is linear (the proof is similar to that in part **(a)**). The matrix M_{T_e} (which can be obtained in the same way as M_{T_a} and M_{T_b}) is

$$M_{T_e} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

(g) The map T_g is linear (the proof is similar to that in part **(a)**, but now you have to use the fact that $(u+v)' = u' + v'$ for two polynomials u and v). The matrix M_{T_g} (which can be obtained in the same way as M_{T_a} and M_{T_b}) is

$$M_{T_g} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

(h) The map T_h is linear (the proof is similar to that in part **(a)**). The matrix M_{T_h} (which can be obtained in the same way as M_{T_a} and M_{T_b}) is

$$M_{T_h} = \begin{pmatrix} 0 & 1 & -1 & 1 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

(i) The map T_i is linear (the proof is similar to that in part **(a)**). The matrix M_{T_i} (which can be obtained in the same way as M_{T_a} and M_{T_b}) is

$$M_{T_i} = \begin{pmatrix} 0 & 0 & 2 & -6 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

[*Remark:* You might notice that the image of T_i lies not just in P_2 but actually in the smaller space P_1 (as witnessed by the last row of M_{T_i} consisting of zeroes). So T_i could be defined as a map $P_3 \to P_1$ instead, and then its matrix would be $\begin{pmatrix} 0 & 0 & 2 & -6 \\ 0 & 0 & 0 & 6 \end{pmatrix}$.]

Exercise 0.3. Find the nullspace of the map T_i from Exercise 0.2 **(i)**.

Solution. For any element $f = a + bx + cx^2 + dx^3$ of P_3, we have

$$T_i(f) = T_i\left(a + bx + cx^2 + dx^3\right) = a\underbrace{T_i(1)}_{=0} + b\underbrace{T_i(x)}_{=0} + c\underbrace{T_i\left(x^2\right)}_{=2} + d\underbrace{T_i\left(x^3\right)}_{=6x-6}$$
$$\text{(since } T_i \text{ is linear)}$$
$$= a \cdot 0 + b \cdot 0 + c \cdot 2 + d \cdot (6x - 6) = 6dx + (2c - 6d).$$

Hence, $T_i(f) = 0$ if and only if $6dx + (2c - 6d) = 0$ (as a polynomial in x), which means that both coefficients $6d$ and $2c - 6d$ are 0, which is equivalent to saying that $c = 0$ and $d = 0$.

Thus, the nullspace of T_i consists of all $f = a + bx + cx^2 + dx^3$ for which $c = 0$ and $d = 0$. In other words, the nullspace of T_i consists of all polynomials f having degree ≤ 1.

[*Remark:* I intended this exercise to be an example of how computing the matrix representation of a linear map helps find the nullspace of said map. I didn't expect that finding the nullspace of T_i was easier than this. Here is how this could have been done using the matrix representation:

From Exercise 0.2 **(i)**, we know that T_i is represented by the matrix

$$M_{T_i} = \begin{pmatrix} 0 & 0 & 2 & -6 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

with respect to the monomial bases $(1, x, x^2, x^3)$ and $(1, x, x^2)$ of P_3 and P_2. Now, let $f \in P_3$. Represent f as $f = a + bx + cx^2 + dx^3$ (so that a, b, c, d are the coordinates of f with respect to the monomial basis $(1, x, x^2, x^3)$). Then, $T_i(f) = \widetilde{a} + \widetilde{b}x + \widetilde{c}x^2$, where

$$\begin{pmatrix} \widetilde{a} \\ \widetilde{b} \\ \widetilde{c} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2 & -6 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

Thus, $T_i(f) = 0$ holds if and only if $\begin{pmatrix} \widetilde{a} \\ \widetilde{b} \\ \widetilde{c} \end{pmatrix} = 0$, that is, if and only if $\begin{pmatrix} 0 & 0 & 2 & -6 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = 0$, that is, if and only if $\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$ lies in the nullspace of the matrix $\begin{pmatrix} 0 & 0 & 2 & -6 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. But the nullspace of the matrix $\begin{pmatrix} 0 & 0 & 2 & -6 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

is easy to find (it is spanned by $\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\\0\\0\end{pmatrix}$), and so we conclude that $T_i(f) = 0$ if and only if $\begin{pmatrix}a\\b\\c\\d\end{pmatrix}$ is a linear combination of $\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\\0\\0\end{pmatrix}$. In other words, $T_i(f) = 0$ if and only if $c = 0$ and $d = 0$. In other words, $T_i(f) = 0$ if and only if f has degree ≤ 1. Hence, we again conclude that the nullspace of T_i consists of all polynomials f having degree ≤ 1.]

Exercise 0.4. The vector space P_n has many other bases. One of them is

$$\left(\binom{x}{0}, \binom{x}{1}, ..., \binom{x}{n}\right),$$

where $\binom{x}{i}$ is defined as the polynomial $\dfrac{x(x-1)\cdots(x-i+1)}{i!}$ (this is the classical definition of binomial coefficients, except that x is now a polynomial variable rather than an integer). This is called the *binomial basis* of P_n.

(a) Represent the maps T_a, T_b, T_g and T_h of Exercise 0.2 with respect to the binomial bases.

(b) Find the general pattern:

(b$_1$) for the matrix representing the map $T_{a,n} : P_n \to P_{n+1}$ given by $T_{a,n}(f) = f(x+1)$;

(b$_2$) for the matrix representing the map $T_{b,n} : P_n \to P_{n+1}$ given by $T_{b,n}(f) = xf(x)$;

(b$_3$) for the matrix representing the map $T_{g,n} : P_n \to P_{n-1}$ given by $T_{g,n}(f) = f'(x)$.

(b$_4$) for the matrix representing the map $T_{h,n} : P_n \to P_{n-1}$ given by $T_{h,n}(f) = f(x) - f(x-1)$.

Solution. **(a)** Let us first see how to change bases between the monomial basis $(1, x, x^2, \ldots, x^n)$ and the binomial basis $\left(\binom{x}{0}, \binom{x}{1}, ..., \binom{x}{n}\right)$ of P_n. There are general ways to do this by inverting matrices, but in this case a faster method exists:

If we are given the expansion of a polynomial $f \in P_n$ in the binomial basis $\left(\binom{x}{0}, \binom{x}{1}, ..., \binom{x}{n}\right)$, then we can obtain its expansion in the monomial basis

$\left(1, x, x^2, \ldots, x^n\right)$ if we know how to expand every $\dbinom{x}{i}$ in the monomial basis. This is easy: we just need to multiply out the product in the definition $\dbinom{x}{i} = \dfrac{x(x-1)\cdots(x-i+1)}{i!}$.

The converse direction is more interesting. How do we expand a polynomial $f \in P_n$ in the binomial basis $\left(\dbinom{x}{0}, \dbinom{x}{1}, ..., \dbinom{x}{n}\right)$? In other words, how do we write f as a linear combination of $\dbinom{x}{0}, \dbinom{x}{1}, ..., \dbinom{x}{n}$?

Let us see an example: Say, $n = 3$ and $f = x^3 - 2x^2 + 4x - 1$. We have $\dbinom{x}{3} = \dfrac{1}{6}x^3 - \dfrac{1}{2}x^2 + \dfrac{1}{3}x$; therefore, $6\dbinom{x}{3} = x^3 - 3x^2 + 2x$ starts with x^3. Hence, subtracting $6\dbinom{x}{3}$ from f yields a polynomial of smaller degree than f:

$$f - 6\binom{x}{3} = x^2 + 2x - 1. \tag{1}$$

Next, we notice that $\dbinom{x}{2} = \dfrac{1}{2}x^2 - \dfrac{1}{2}x$, whence $2\dbinom{x}{2} = x^2 - x$, which begins with an x^2 just as $x^2 + 2x - 1$ does. Hence, subtracting the equality $2\dbinom{x}{2} = x^2 - x$ from the equality (1) yields

$$f - 6\binom{x}{3} - 2\binom{x}{2} = 3x - 1.$$

Next, we subtract $3\dbinom{x}{1} = 3x$ from this equality, thus kicking out the $3x$ term:

$$f - 6\binom{x}{3} - 2\binom{x}{2} - 3\binom{x}{1} = -1.$$

Finally, we get rid of the -1 on the right hand side by subtracting $-\dbinom{x}{0} = -1$:

$$f - 6\binom{x}{3} - 2\binom{x}{2} - 3\binom{x}{1} - \binom{x}{0} = 0.$$

Thus,

$$f = 6\binom{x}{3} + 2\binom{x}{2} + 3\binom{x}{1} + \binom{x}{0}.$$

Thus, we have expanded our f as a linear combination of $\dbinom{x}{0}, \dbinom{x}{1}, ..., \dbinom{x}{n}$.

The method we followed here works in all generality: If we want to expand a degree-d polynomial f as a linear combination of $\binom{x}{0}, \binom{x}{1}, ..., \binom{x}{n}$, we subtract an appropriate scalar multiple of $\binom{x}{d}$ from it to get rid of its x^d-term, and then are left with a polynomial of degree smaller than d; we then proceed in the same way until we clear out the last term and end up with 0. Then, f is the sum of all multiples of $\binom{x}{d}$'s subtracted.

We can now get our hands dirty and find the matrix representation M_{T_a} of T_a with respect to the binomial bases $\left(\binom{x}{0}, \binom{x}{1}, \binom{x}{2}\right)$ and $\left(\binom{x}{0}, \binom{x}{1}, \binom{x}{2}\right)$ of P_2 and P_2 using part **2.** of the repetitorium. We have

$$T_a\left(\binom{x}{0}\right) = \binom{x}{0} |_{x+1} = \binom{x+1}{0} = 1 = 1 \cdot \binom{x}{0} + 0 \cdot \binom{x}{1} + 0 \cdot \binom{x}{2} + 0 \cdot \binom{x}{3};$$
$$T_a\left(\binom{x}{1}\right) = \binom{x}{1} |_{x+1} = \binom{x+1}{1} = x+1 = 1 \cdot \binom{x}{0} + 1 \cdot \binom{x}{1} + 0 \cdot \binom{x}{2} + 0 \cdot \binom{x}{3};$$
$$T_a\left(\binom{x}{2}\right) = \binom{x}{2} |_{x+1} = \binom{x+1}{2} = \frac{x(x+1)}{2} = 0 \cdot \binom{x}{0} + 1 \cdot \binom{x}{1} + 1 \cdot \binom{x}{2} + 0 \cdot \binom{x}{3}.$$

Thus, the matrix M_{T_a} is

$$M_{T_a} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Similar computations show that

$$M_{T_b} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix};$$
$$M_{T_g} = \begin{pmatrix} 0 & 1 & -1/2 & 1/3 \\ 0 & 0 & 1 & -1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix};$$
$$M_{T_h} = \begin{pmatrix} 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(b) This is a tricky problem; don't expect to see it on the exam. But things like this are what mathematics feels like. If you only want to prepare for the exam, scroll down to the next exercise.

(b$_1$) The matrix $M_{T_{a,n}}$ representing the map $T_{a,n} : P_n \to P_{n+1}$ with respect to the binomial bases $\left(\binom{x}{0}, \binom{x}{1}, ..., \binom{x}{n}\right)$ of P_n and $\left(\binom{x}{0}, \binom{x}{1}, ..., \binom{x}{n+1}\right)$ of

P_{n+1} is

$$M_{T_{a,n}} = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

which is the matrix with 1's along the main diagonal, 1's along the diagonal above, and 0's everywhere else.

To see why this is so, we only need to check that

$$T_{a,n}\left(\binom{x}{k}\right) = \binom{x}{k} + \binom{x}{k-1} \qquad \text{for every } k \in \{1, 2, \ldots, n\} \tag{2}$$

and that

$$T_{a,n}\left(\binom{x}{0}\right) = \binom{x}{0}. \tag{3}$$

Checking (3) is trivial (remember that $\binom{x}{0} = 1$), so let us check (2).

Given $k \in \{1, 2, \ldots, n\}$, we have

$$\begin{aligned}
T_{a,n}\left(\binom{x}{k}\right) &= \binom{x}{k} \mid_{x+1} \qquad \text{(by the definition of } T_{a,n}\text{)} \\
&= \frac{x(x-1)\cdots(x-k+1)}{k!} \mid_{x+1} \qquad \left(\text{by the definition of } \binom{x}{k}\right) \\
&= \frac{(x+1)((x+1)-1)\cdots((x+1)-k+1)}{k!} \\
&= \frac{(x+1)x\cdots(x-k+2)}{k!} = \underbrace{(x+1)}_{=((x-k+1)+k)} \cdot \frac{x(x-1)\cdots(x-k+2)}{k!} \\
&= ((x-k+1)+k) \cdot \frac{x(x-1)\cdots(x-k+2)}{k!} \\
&= \underbrace{(x-k+1) \cdot \frac{x(x-1)\cdots(x-k+2)}{k!}}_{=\frac{x(x-1)\cdots(x-k+2)(x-k+1)}{k!}=\binom{x}{k}} \\
&\quad + \underbrace{k \cdot \frac{x(x-1)\cdots(x-k+2)}{k!}}_{=\frac{x(x-1)\cdots(x-k+2)}{k!/k}\underset{(\text{since } k!/k=(k-1)!)}{=}\frac{x(x-1)\cdots(x-k+2)}{(k-1)!}} \\
&= \binom{x}{k} + \underbrace{\frac{x(x-1)\cdots(x-k+2)}{(k-1)!}}_{=\binom{x}{k-1}} = \binom{x}{k} + \binom{x}{k-1}.
\end{aligned}$$

This was actually the boring way to prove (2). Here is a more interesting one (a bit of overkill for this problem, but the idea shows its power eventually):

Fix $k \in \{1, 2, \ldots, n\}$. We need to prove the identity $T_{a,n}\left(\binom{x}{k}\right) = \binom{x}{k} + \binom{x}{k-1}$. In other words, we need to prove the identity $\binom{x+1}{k} = \binom{x}{k} + \binom{x}{k-1}$ (since $T_{a,n}\left(\binom{x}{k}\right) = \binom{x}{k} \mid_{x+1} = \binom{x+1}{k}$). This is an identity between polynomials in x. But two polynomials are identical if they are equal at infinitely many points (because a polynomial of degree d is uniquely determined by $d+1$ values); in particular, they are identical if they are equal at all positive integers. Thus, in order to prove the identity (2) between polynomials, it is enough to verify the identity $\binom{x+1}{k} = \binom{x}{k} + \binom{x}{k-1}$, it is enough to prove the equality

$$\binom{N+1}{k} = \binom{N}{k} + \binom{N}{k-1} \qquad \text{for every positive integer } N. \tag{4}$$

How do we prove (4) ? Fix a positive integer N. The number of all k-element subsets of $\{1, 2, \ldots, N+1\}$ is known to be $\binom{N+1}{k}$ (according to the combinatorial definition of binomial coefficients). But on the other hand, one can choose a k-element subset of $\{1, 2, \ldots, N+1\}$ by first deciding whether $N+1$ goes into the subset or not, and then

- if $N+1$ does not go into the subset, choosing the remaining k elements from $\{1, 2, \ldots, N\}$ in one of $\binom{N}{k}$ possible ways.
- if $N+1$ goes into the subset, choosing the remaining $k-1$ elements from $\{1, 2, \ldots, N\}$ in one of $\binom{N}{k-1}$ possible ways.

Hence, the total number of k-element subsets of $\{1, 2, \ldots, N+1\}$ is $\binom{N}{k} + \binom{N}{k-1}$. Contrasting this number with the $\binom{N+1}{k}$ we have obtained before, we obtain $\binom{N+1}{k} = \binom{N}{k} + \binom{N}{k-1}$. Thus, (4) is proven, and with it the identity $\binom{x+1}{k} = \binom{x}{k} + \binom{x}{k-1}$, and thus (2).

Generally, when one works with binomial coefficients, combinatorics is never far away. We have not been able to interpret the **polynomials** $\binom{x}{k}$ combinatorially, but we know how to interpret their **values** $\binom{N}{k}$ at positive N (actually, at nonnegative N if you are pedantic) combinatorially, and this was enough for our proof because a polynomial is uniquely determined by infinitely many values.

($\mathbf{b_2}$) The matrix $M_{T_{b,n}}$ representing the map $T_{b,n} : P_n \to P_{n+1}$ with respect to the binomial bases $\left(\binom{x}{0}, \binom{x}{1}, ..., \binom{x}{n}\right)$ of P_n and $\left(\binom{x}{0}, \binom{x}{1}, ..., \binom{x}{n+1}\right)$ of P_{n+1} is

$$M_{T_{b,n}} = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 2 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 3 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & n+1 \end{pmatrix}.$$

This is the matrix whose j-th column (for each j) has a $j-1$ in its j-th entry and a j in its $(j+1)$-st entry.

To see why this is so, it suffices to prove that

$$T_{b,n}\left(\binom{x}{k}\right) = k\binom{x}{k} + (k+1)\binom{x}{k+1} \qquad \text{for every } k \in \{0,1,\ldots,n\}. \quad (5)$$

This has a simple algebraic proof, which begins by noticing that

$$\underbrace{T_{b,n}\left(\binom{x}{k}\right)}_{=x\binom{x}{k}} - k\binom{x}{k} = x\binom{x}{k} - k\binom{x}{k} = (x-k)\binom{x}{k}$$

and rewriting this further. I am leaving this to you.

(b$_3$) The matrix $M_{T_{g,n}}$ representing the map $T_{g,n} : P_n \to P_{n-1}$ with respect to the binomial bases $\left(\binom{x}{0}, \binom{x}{1}, \ldots, \binom{x}{n}\right)$ of P_n and $\left(\binom{x}{0}, \binom{x}{1}, \ldots, \binom{x}{n-1}\right)$ of P_{n-1} is

$$M_{T_{g,n}} = \begin{pmatrix} 0 & 1 & -1/2 & 1/3 & \cdots & (-1)^{n-2}/(n-1) & (-1)^{n-1}/n \\ 0 & 0 & 1 & -1/2 & \cdots & (-1)^{n-3}/(n-2) & (-1)^{n-2}/(n-1) \\ 0 & 0 & 0 & 1 & \cdots & (-1)^{n-4}/(n-3) & (-1)^{n-3}/(n-2) \\ 0 & 0 & 0 & 0 & \cdots & (-1)^{n-5}/(n-4) & (-1)^{n-4}/(n-3) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

The proof of this boils down to the identity

$$T_{g,n}\left(\binom{x}{k}\right) = -\sum_{j=1}^{k} \frac{(-1)^j}{j}\binom{x}{k-j} \qquad \text{for every } k \in \{1,2,\ldots,n\}.$$

This is harder to prove than the preceding ones.[4]

(b$_4$) The matrix $M_{T_{h,n}}$ representing the map $T_{h,n} : P_n \to P_{n-1}$ with respect to the binomial bases $\left(\binom{x}{0}, \binom{x}{1}, \ldots, \binom{x}{n}\right)$ of P_n and $\left(\binom{x}{0}, \binom{x}{1}, \ldots, \binom{x}{n-1}\right)$ of

[4]For the curious with some familiarity with generating functions: differentiate $(1+t)^x = \exp(x\log(1+t))$ (an identity between power series in two variables x and t) with respect to x.

P_{n-1} is

$$M_{T_{g,n}} = \begin{pmatrix} 0 & 1 & -1 & 1 & \cdots & (-1)^{n-2} & (-1)^{n-1} \\ 0 & 0 & 1 & -1 & \cdots & (-1)^{n-3} & (-1)^{n-2} \\ 0 & 0 & 0 & 1 & \cdots & (-1)^{n-4} & (-1)^{n-3} \\ 0 & 0 & 0 & 0 & \cdots & (-1)^{n-5} & (-1)^{n-4} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

The proof of this boils down to the identity

$$T_{h,n}\left(\binom{x}{k}\right) = -\sum_{j=1}^{k} (-1)^j \binom{x}{k-j} \qquad \text{for every } k \in \{1, 2, \ldots, n\}.$$

This can be proven by induction over k.

Exercise 0.5. The *total degree* of a monomial in several variables is the sum of the exponents of the variables. For example, the total degree of x^2y^5 is $2+5=7$.

The *total degree* of a polynomial in several variables is the maximal total degree of any of its monomials. For instance, the total degree of $2x^2y - 3xy^2 + 5x^2y^2 + 6x^3y - 7x$ is 4.

Given $n \in \mathbb{N}$, the polynomials in two variables x and y having total degree $\leq n$ form a vector space. What is the dimension of this space?

[**Example:** For $n = 2$, this space consists of all polynomials of the form $ax^2 + bxy + cy^2 + dx + ey + f$.]

Solution. The vector space of all polynomials in two variables x and y having total degree $\leq n$ has a basis consisting of all monomials of total degree $\leq n$. Thus, its dimension is the number of all such monomials. What is this number?

The monomials of total degree $\leq n$ are the monomials of the form x^iy^j with $i + j \leq n$. These are the monomials

$$\begin{array}{lllll} x^0y^0, & x^0y^1, & \ldots, & x^0y^{n-1}, & x^0y^n, \\ x^1y^0, & x^1y^1, & \ldots, & x^1y^{n-1}, & \\ \cdots & & & & \\ x^{n-1}y^0, & x^{n-1}y^1, & & & \\ x^ny^0 & & & & \end{array},$$

and their total number is $\dfrac{(n+1)(n+2)}{2}$ (in fact, for every fixed i, there are $n-i+1$ of them, and so the total number is $(n+1) + n + (n-1) + \cdots + 1 = \dfrac{(n+1)(n+2)}{2}$). So the dimension is $\dfrac{(n+1)(n+2)}{2}$.

For fixed integers n and m, the $m \times n$-matrices themselves form a vector space: you can add two $m \times n$-matrices, you can multiply an $m \times n$-matrix with a scalar, and there is a zero $m \times n$-matrix. This space is mn-dimensional.

Exercise 0.6. For every $n \in \mathbb{N}$, let M_n be the vector space of all $n \times n$-matrices. Which of the following maps is linear?

(a) The map $M_2 \to P_2$ sending every matrix $A \in M_2$ to the characteristic polynomial of A ?

(b) The map $M_2 \to \mathbb{R}$ sending every matrix $A \in M_2$ to $\det A$?

(c) The map $M_2 \to \mathbb{R}$ sending every matrix $A \in M_2$ to $\operatorname{Tr} A$?

Solution. **(a)** No. This map sends the zero matrix $0 \in M_2$ to the polynomial λ^2, which is not zero. But a linear map would have to send 0 to 0.

(b) No. It fails the second axiom for a linear map (see part **1.** of the repetitorium): it does not satisfy $\det(u+v) = \det u + \det v$ for $u, v \in M_2$. For an example, take $u = I_2$ and $v = I_2$; then $\det(u+v) = \det(I_2 + I_2) = \det(2I_2) = 4$, which differs from $\det u + \det v = \det I_2 + \det I_2 = 1 + 1 = 2$.

(c) Yes. To prove this, we need to check the three axioms of a linear map. We will only show the second one, since the other two are similar.

So we need to prove that $\operatorname{Tr}(u+v) = \operatorname{Tr} u + \operatorname{Tr} v$ for any $u, v \in M_2$. (Note that u and v are matrices here. I only denote them by lowercase letters to match the notation in the repetitorium.)

Write $u = \begin{pmatrix} u_{1,1} & u_{1,2} \\ u_{2,1} & u_{2,2} \end{pmatrix}$ and $v = \begin{pmatrix} v_{1,1} & v_{1,2} \\ v_{2,1} & v_{2,2} \end{pmatrix}$. Then, $u + v = \begin{pmatrix} u_{1,1} + v_{1,1} & u_{1,2} + v_{1,2} \\ u_{2,1} + v_{2,1} & u_{2,2} + v_{2,2} \end{pmatrix}$, so that

$$\operatorname{Tr}(u+v) = (u_{1,1} + v_{1,1}) + (u_{2,2} + v_{2,2}) = \underbrace{(u_{1,1} + u_{2,2})}_{=\operatorname{Tr} u} + \underbrace{(v_{1,1} + v_{2,2})}_{=\operatorname{Tr} v} = \operatorname{Tr} u + \operatorname{Tr} v,$$

which is exactly what we need to prove.

[Of course, this is not specific to 2×2-matrices. For every $n \in \mathbb{N}$, the map $M_n \to \mathbb{R}$ sending every $n \times n$-matrix A to $\operatorname{Tr} A$ is linear.]

0.2. Positive definiteness

0.2.1. Repetitorium

A symmetric $n \times n$-matrix A is positive definite if and only if it satisfies any of the following **equivalent** criteria:

- Every nonzero vector $x \in \mathbb{R}^n$ satisfies $x^T A x > 0$. (This is what Strang calls the "energy-based" definition.)
- All eigenvalues of A are positive. (They are real because A is symmetric.)
- All n upper-left determinants of A are positive. (Recall that the k-th *upper-left determinant* of A (for $1 \leq k \leq n$) is the determinant of the matrix obtained by taking only the first k rows and the first k columns of A. For in-

stance, the 2-nd upper-left determinant of $\begin{pmatrix} 1 & 2 & 4 \\ 2 & 6 & 4 \\ 4 & 4 & 3 \end{pmatrix}$ is $\det \begin{pmatrix} 1 & 2 \\ 2 & 6 \end{pmatrix} = 2$. The n-th upper-left determinant of A is $\det A$.)

- Gaussian elimination on A can be performed without switching rows, and all n pivots are positive.

- The matrix A can be written in the form $R^T R$ for a matrix R whose columns are linearly independent (i.e., whose nullspace is 0).

- The matrix A can be written in the form $R^T DR$ for a matrix B whose columns are linearly independent and a diagonal matrix D with positive diagonal entries.

Depending on the problem at hand, you might find one or another of these criteria easier to check. The eigenvalues test is usually unsuitable because it requires you to compute the eigenvalues!

Keep in mind that the matrix has to be *symmetric* for us to speak of it being positive definite. $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ does not count as positive definite even though we do have $x^T Ax > 0$ for every nonzero $x \in \mathbb{R}^n$, for lack of symmetry.

Example: The symmetric matrix $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ is positive definite, for any of the following reasons:

- Every nonzero vector $x \in \mathbb{R}^2$ satisfies $x^T Ax > 0$. This is because if we write x as $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, then $x^T Ax = x_1^2 + 2x_2^2 + 2x_1 x_2 = (x_1 + x_2)^2 + x_2^2$, which is > 0 unless $x = 0$.

- All eigenvalues of A are positive, because they are $\frac{1}{2}\sqrt{5} + \frac{3}{2}$ and $\frac{3}{2} - \frac{1}{2}\sqrt{5}$.

- The upper-left determinants of A are positive, being $\det(1) = 1$ and $\det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = 1$.

- Gaussian elimination of A works without switching rows and gives rise to the pivots 1 and 1, which are manifestly positive.

0.2.2. Exercises

Exercise 0.7. If a symmetric 3×3-matrix $\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$ is positive definite, then show that $df > e^2$.

First solution. Let $\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$ be a positive definite symmetric 3×3-matrix.

The "energy-based" definition of positive definiteness then yields that $x^T \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} x > 0$ for all nonzero $x \in \mathbb{R}^3$. Consequently, $\begin{pmatrix} 0 \\ 1 \\ y \end{pmatrix}^T \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ y \end{pmatrix} > 0$ for all $y \in \mathbb{R}$. But this rewrites as $fy^2 + 2ey + d > 0$. The quadratic $fy^2 + 2ey + d$ (as a function of y) thus never meets the x-axis. Therefore, its discriminant $(2e)^2 - 4fd$ is negative. In other words, $(2e)^2 - 4fd < 0$. This quickly yields $df > e^2$.

Second solution. Here is a solution which is probably more motivated.

Let $A = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$ be a positive definite symmetric 3×3-matrix. We shall show that the matrix $\widetilde{A} = \begin{pmatrix} f & e & c \\ e & d & b \\ c & b & a \end{pmatrix}$ obtained from A by "180°-rotation around its center" is also positive definite. Once this is shown, it will follow that the upper-left determinants of $\widetilde{A}$ are positive (due to one of the criteria for positive definite). Since one of these determinants is $fd - ee = df - e^2$, this will yield that $df - e^2 > 0$, so that $df > e^2$, and the problem will be solved.

It remains to show that $\widetilde{A}$ is positive definite.

The "energy-based" definition of positive definiteness then yields that $x^T A x > 0$ for all nonzero $x \in \mathbb{R}^3$ (since A is positive definite). In other words, $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^T A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} > 0$ for all nonzero $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$. In other words, $\begin{pmatrix} x_3 \\ x_2 \\ x_1 \end{pmatrix}^T \widetilde{A} \begin{pmatrix} x_3 \\ x_2 \\ x_1 \end{pmatrix} > 0$ for all nonzero $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$ (because an easy computation confirms that $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^T A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_3 \\ x_2 \\ x_1 \end{pmatrix}^T \widetilde{A} \begin{pmatrix} x_3 \\ x_2 \\ x_1 \end{pmatrix}$). In yet other words, $x^T \widetilde{A} x > 0$ for all nonzero $x \in \mathbb{R}^3$ (now, we have substituted x for $\begin{pmatrix} x_3 \\ x_2 \\ x_1 \end{pmatrix}$). This shows that $\widetilde{A}$ is positive definite (by the "energy definition"), and we are done.

Exercise 0.8. What entries of a positive definite symmetric matrix can be 0 ? (For instance, can the $(2,2)$-nd entry be 0 ? Can the $(2,3)$-th entry?)

Solution. Every off-diagonal entry can be 0 (because the $n \times n$ identity matrix is a positive definite matrix and has all its off-diagonal entries equal to 0).

On the other hand, diagonal entries cannot be 0. Let us see why. Assume that a positive definite symmetric $n \times n$-matrix A has its (i,i)-th entry equal to 0. But this entry is precisely $e_i^T A e_i$, where e_i is the i-th standard basis vector (that is, the column vector with a 1 in its i-th row and 0's everywhere else). So $e_i^T A e_i = 0$. This contradicts the fact that $x^T A x > 0$ for every nonzero $x \in \mathbb{R}^n$ (because A is positive definite). So our assumption was wrong, and thus diagonal entries of a positive definite symmetric matrix cannot be zero. (Actually, this argument shows that they must be positive.)

Exercise 0.9. If A is a positive definite symmetric $n \times n$-matrix, and B is an $n \times m$-matrix, then:

(a) If the columns of B are linearly independent, then show that $B^T A B$ is positive definite.

(b) If the columns of B are **not** linearly independent, then show that $B^T A B$ is **not** positive definite.

Solution. Let us first notice that $A^T = A$ (since A is symmetric) and thus $\left(B^T A B\right)^T = B^T \underbrace{A^T}_{=A} \underbrace{\left(B^T\right)^T}_{=B} = B^T A B$. Hence, $B^T A B$ is symmetric.

(a) We know that A is positive definite. Thus,

$$y^T A y > 0 \tag{6}$$

for every nonzero $y \in \mathbb{R}^n$ (by the "energy-based" definition of positive definiteness).

The columns of B are linearly independent. In other words, the nullspace of B is 0.

Now, let $x \in \mathbb{R}^m$ be nonzero. Then, $Bx \neq 0$ (since otherwise, x would lie in the nullspace of B, but this nullspace is 0). Thus, (6) (applied to $y = Bx$) yields $(Bx)^T ABx > 0$. Since $\underbrace{(Bx)^T}_{=x^T B^T} ABx = x^T B^T ABx = x^T \left(B^T AB\right) x$, this rewrites as $x^T \left(B^T AB\right) x > 0$.

So we have shown that $x^T \left(B^T AB\right) x > 0$ for every nonzero $x \in \mathbb{R}^m$. Combined with the symmetry of $B^T AB$, this yields that $B^T AB$ is positive definite.

(b) The columns of B are not linearly independent. In other words, the nullspace of B is **not** 0. So there exists a nonzero $x \in \mathbb{R}^m$ such that $Bx = 0$. For this x, we must have $x^T \left(B^T AB\right) x = x^T B^T A \underbrace{Bx}_{=0} = 0$. This means that not every nonzero $x \in \mathbb{R}^m$ can satisfy $x^T \left(B^T AB\right) x > 0$. Thus, $B^T AB$ is not positive definite.

[But a reasoning similar to that for part **(a)** shows that $B^T AB$ is positive semidefinite.]

Exercise 0.10. For which real numbers p is the matrix $\begin{pmatrix} 1 & p & 0 \\ p & 1 & 1-p \\ 0 & 1-p & 1 \end{pmatrix}$ positive definite?

Solution. The matrix $\begin{pmatrix} 1 & p & 0 \\ p & 1 & 1-p \\ 0 & 1-p & 1 \end{pmatrix}$ is always a symmetric 3×3-matrix. Hence, by one of our criteria, we know that it is positive definite if and only if all 3 of its upper-left determinants are positive. These determinants are $\det(1) = 1$, $\det \begin{pmatrix} 1 & p \\ p & 1 \end{pmatrix} = 1 - p^2$ and $\det \begin{pmatrix} 1 & p & 0 \\ p & 1 & 1-p \\ 0 & 1-p & 1 \end{pmatrix} = 2p - 2p^2$. Thus, our matrix is positive definite if and only if all of 1, $1 - p^2$ and $2p - 2p^2$ are positive definite. This is easily seen to hold if and only if $0 < p < 1$.

Exercise 0.11. Given an integer $n \geq 2$. For which real numbers p is the $n \times n$-matrix $\begin{pmatrix} p & 1 & \cdots & 1 \\ 1 & p & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & p \end{pmatrix}$ positive definite?

Solution. For $p > 1$.

Let us prove this. We denote the matrix $\begin{pmatrix} p & 1 & \cdots & 1 \\ 1 & p & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & p \end{pmatrix}$ by A. Clearly, A is symmetric.

We shall first show that A is positive definite whenever $p > 1$. Indeed, assume

$p > 1$. Let $x \in \mathbb{R}^n$ be nonzero. Write x in the form $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. Then,

$$\begin{aligned} x^T A x &= p x_1^2 + p x_2^2 + \cdots + p x_n^2 \\ &\qquad + 2 x_1 x_2 + 2 x_1 x_3 + \cdots + 2 x_1 x_n \\ &\qquad + 2 x_2 x_3 + 2 x_2 x_4 + \cdots + 2 x_2 x_n \\ &\qquad + \cdots \\ &\qquad + 2 x_{n-1} x_n \\ &= x_1^2 + x_2^2 + \cdots + x_n^2 \\ &\qquad + 2 x_1 x_2 + 2 x_1 x_3 + \cdots + 2 x_1 x_n \\ &\qquad + 2 x_2 x_3 + 2 x_2 x_4 + \cdots + 2 x_2 x_n \\ &\qquad + \cdots \\ &\qquad + 2 x_{n-1} x_n \\ &\qquad + (p-1) \left(x_1^2 + x_2^2 + \cdots + x_n^2 \right) \\ &= (x_1 + x_2 + \cdots + x_n)^2 + (p-1) \left(x_1^2 + x_2^2 + \cdots + x_n^2 \right), \end{aligned}$$

which is manifestly positive (since $p > 1$ and since not all of $x_1, x_2, \ldots, x_n$ are zero). So A is positive definite (according to the "energy-based definition" of positive definiteness).

We now need to show that A is not positive definite unless $p > 1$. Indeed, assume that $p \leq 1$. Let $x = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n$. Then, an easy computation shows that $x^T A x = 2(p-1)$. When $p \leq 1$, this is ≤ 0. Thus, we have found a nonzero $x \in \mathbb{R}^n$ for which $x^T A x \leq 0$. According to the "energy-based definition", this shows that A cannot be positive definite. We are done.