Question: Can we do a similar thing in the polynomial space? Consider bases $\mathcal{B}$ and $\mathcal{C}$ that is not the simple monomial basis:

$$\begin{aligned} \mathcal{B} &= (\underbrace{2\mathbf{m}_2 - 1\mathbf{m}_1}_{=:\mathbf{b}_1}, \underbrace{-8\mathbf{m}_1 - 2\mathbf{m}_0}_{=:\mathbf{b}_2}, \underbrace{1\mathbf{m}_2 + 4\mathbf{m}_1 + 1\mathbf{m}_0}_{=:\mathbf{b}_3}) \\ \text{and} \qquad \mathcal{C} &= (\underbrace{1\mathbf{m}_1 + 1\mathbf{m}_0}_{=:\mathbf{c}_1}, \underbrace{2\mathbf{m}_2 + 2\mathbf{m}_1}_{=:\mathbf{c}_2}, \underbrace{1\mathbf{m}_2 + 1\mathbf{m}_0}_{=:\mathbf{c}_3}). \end{aligned}$$

Answer: Yes, we can do the same by adding the the monomial basis (or a other well-known basis) in the middle. We call the monomial basis by $\mathcal{A}$, which means $\mathcal{A} = (\mathbf{m}_2, \mathbf{m}_1, \mathbf{m}_0)$. Then $T_{\mathcal{A}\leftarrow\mathcal{B}}$ and $T_{\mathcal{A}\leftarrow\mathcal{C}}$ are immediately given:

$$T_{\mathcal{A}\leftarrow\mathcal{B}} = \begin{pmatrix} 2 & 0 & 1 \\ -1 & -8 & 4 \\ 0 & -2 & 1 \end{pmatrix} \qquad \text{and} \qquad T_{\mathcal{A}\leftarrow\mathcal{C}} = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

and then we get $T_{\mathcal{B}\leftarrow\mathcal{C}}$:

$T_{\mathcal{B}\leftarrow\mathcal{C}}$ by using an additional "nice" basis $\mathcal{A}$

$$T_{\mathcal{A}\leftarrow\mathcal{C}} = T_{\mathcal{A}\leftarrow\mathcal{B}} T_{\mathcal{B}\leftarrow\mathcal{C}}$$

$$\Phi_{\mathcal{A}}(\mathbf{x}) \xleftarrow{T_{\mathcal{A}\leftarrow\mathcal{B}}} \Phi_{\mathcal{B}}(\mathbf{x}) \xleftarrow{T_{\mathcal{B}\leftarrow\mathcal{C}}} \Phi_{\mathcal{C}}(\mathbf{x})$$

and hence $\quad T_{\mathcal{B}\leftarrow\mathcal{C}} = (T_{\mathcal{A}\leftarrow\mathcal{B}})^{-1} T_{\mathcal{A}\leftarrow\mathcal{C}}.$

Since we again have to find an inverse of a matrix, we can use the Gauß-Jordan algorithm again:

$$(T_{\mathcal{A}\leftarrow\mathcal{B}} \mid T_{\mathcal{A}\leftarrow\mathcal{C}}) \rightsquigarrow (\mathbb{1} \mid T_{\mathcal{B}\leftarrow\mathcal{C}}). \tag{7.12}$$

For our example, this gives us:

$$\left(\begin{array}{ccc|ccc} 2 & 0 & 1 & 0 & 2 & 1 \\ -1 & -8 & 4 & 1 & 2 & 0 \\ 0 & -2 & 1 & 1 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \;\middle|\; \boxed{\begin{matrix} 3 & -2 & 4 \\ -7/2 & 3 & -4 \\ -6 & 6 & -7 \end{matrix}} \right).$$

VL19 ↓

The boxed matrix is indeed $T_{\mathcal{B}\leftarrow\mathcal{C}}$.

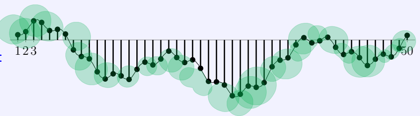
V abstract vector space $\qquad \mathcal{B} = (b_1, \dots, b_n)$

each $v \in V \rightsquigarrow \Phi_{\mathcal{B}}(v) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{F}^n$, $\quad v = \alpha_1 b_1 + \dots + \alpha_n b_n$

What is a good basis for your problem?

Change of basis for audio: WAV vs. MP3

Assume you have an audio signal $\mathbf{f}$ *given at finite time steps* $t = 1, 2, ..., 50$ *(e.g. milliseconds). Hence, you have some measure values* $f_1, f_2, \ldots, f_{50} \in \mathbb{R}$.

$\mathbf{f} =$ 

$\mathbf{b}_1$: 

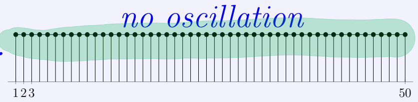
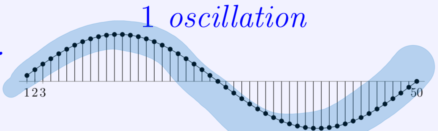
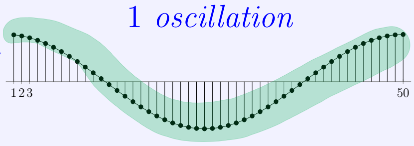
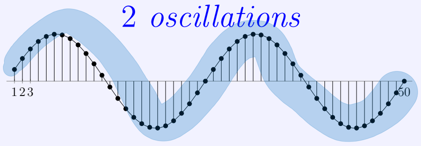
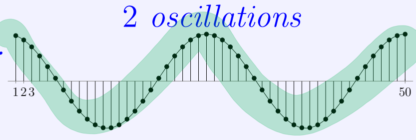
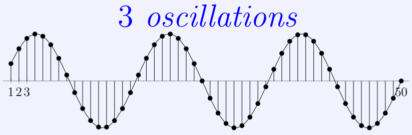
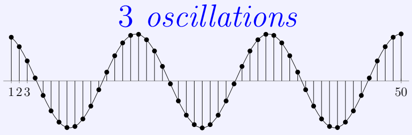
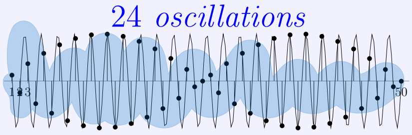
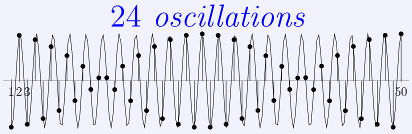
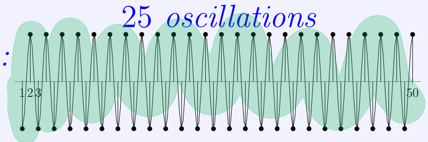
$\mathbf{b}_2$: 

$\vdots$

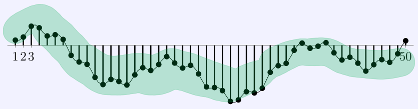
$\mathbf{b}_{50}$: 

At this point you know that the audio signal $\mathbf{f}$ *is a vector in a 50-dimensional space, which can be represented with respect to the canonical basis* $\mathcal{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_{50})$ *of* $\mathbb{R}^{50}$. *The coordinates of* $\mathbf{f}$ *are exactly the values* $f_1, f_2, \ldots, f_{50}$. *For describing tones (so oscillations) this basis is not optimal!*

We want to change to a basis $\mathcal{C}$ *of* $\mathbb{R}^{50}$*, which is better fitting for tones.*

Sine waves	***Cosine waves***
	no oscillation $\mathbf{c}_{25}$: 
1 oscillation $\mathbf{c}_1$: 	*1 oscillation* $\mathbf{c}_{26}$: 
2 oscillations $\mathbf{c}_2$: 	*2 oscillations* $\mathbf{c}_{27}$: 
3 oscillations $\mathbf{c}_3$: 	*3 oscillations* $\mathbf{c}_{28}$: 
$\vdots$	$\vdots$
24 oscillations $\mathbf{c}_{24}$: 	*24 oscillations* $\mathbf{c}_{49}$: 
	25 oscillations $\mathbf{c}_{50}$: 

One can show: $\mathcal{C} = (\mathbf{c}_1, \dots, \mathbf{c}_{50})$ *is also linearly independent and hence a basis of* $\mathbb{R}^{50}$.

$\mathbf{f} =$ 

$$\Phi_{\mathcal{C}}(f) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ -1 \\ 2 \\ \vdots \\ 1/4 \\ \vdots \end{pmatrix}$$

The signal $\mathbf{f}$ *from above has the following form:*

$$\mathbf{f} = \mathbf{c}_3 - \mathbf{c}_{25} + 3\mathbf{c}_{26} + \tfrac{1}{4}\mathbf{c}_{40}.$$

We reckon that most signals $\mathbf{f}$ *are a superposition of some "basic tones"* $\mathbf{c}_i$.

Compression: One stores only the coordinates in $\Phi_{\mathcal{C}}(\mathbf{f})$*. One can also focus on the (for humans) important frequencies and ignore the higher and lower ones (e.g. MP3 file format). All this saves storage space instead of storing the coordinates* $\Phi_{\mathcal{B}}(\mathbf{f}) = (f_1, \dots, f_{50})^T$ *(e.g. WAV file format). Similar ideas exist for two-dimensional signals like pictures:* $\Rightarrow$ *BMP vs. JPG.*

$\Phi_{\mathcal{B}}$

Picture $\in \mathbb{R}^{50 \times 50}$

Information: The change of basis from $\mathcal{B}$ *to* $\mathcal{C}$ *is important for a lot of applications and known as the Fourier transform. We will consider it in more detail in the analysis lecture.*

7.5 General vector space with inner product and norms

Recall that in the vector spaces $\mathbb{R}^n$ and $\mathbb{C}^n$, besides the algebraic structure given by

vector addition $+$ and the scalar multiplication $\cdot$,

we also defined a geometric structure by choosing

an inner product $\langle \cdot, \cdot \rangle$ and also a norm $\|\cdot\|$

for measuring angles and lengths.

Attention! Convention for $\mathbb{F}=\mathbb{R}$ and $\mathbb{F}=\mathbb{C}$

Since we handle the cases $\mathbb{F} = \mathbb{R}$ *and* $\mathbb{F} = \mathbb{C}$ *simultaneously, we also use the notion of the complex conjugation in the real case. Hence, for* $\alpha \in \mathbb{F}$ *we write:*

$$\overline{\alpha} := \begin{cases} \alpha & \text{if } \mathbb{F} = \mathbb{R}, \\ \overline{\alpha} & \text{if } \mathbb{F} = \mathbb{C} \quad \text{(complex conjugate number).} \end{cases}$$

Analogously, for a matrix $A \in \mathbb{F}^{m \times n}$ *with* $m, n \in \mathbb{N}$:

$$A^* := \begin{cases} A^T & \text{if } \mathbb{F} = \mathbb{R} \quad \text{(transpose)}, \\ A^* & \text{if } \mathbb{F} = \mathbb{C} \quad \text{(adjoint).} \end{cases}$$

7.5.1 Inner products

Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and V be an $\mathbb{F}$-vector space.

Definition 7.23. Inner product

A map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ is called an inner product for V if it fulfils: For all $\mathbf{x}, \mathbf{x}', \mathbf{y} \in V$ and $\alpha \in \mathbb{F}$:

(S1) $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ *for all* $\mathbf{x} \neq \mathbf{o}$, *(positive definite)*

(S2) $\langle \mathbf{x} + \mathbf{x}', \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}', \mathbf{y} \rangle$, *(additive)*

(S3) $\langle \alpha \mathbf{x}, \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle$, *(homogeneous)*

$\Big\}$ *(linear)*

(S4) $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$. *((conj.) symmetric)*

A vector space with an inner product is often called a pre-Hilbert space.

Recall all the properties we could derive from these four rules. For example:

$$\langle \mathbf{x}, \alpha \mathbf{y} \rangle = \overline{\alpha} \langle \mathbf{x}, \mathbf{y} \rangle \qquad \text{for all} \quad \alpha \in \mathbb{F}, \ \mathbf{x}, \mathbf{y} \in V.$$

The proof goes like: $\langle \mathbf{x}, \alpha \mathbf{y} \rangle \overset{(S4)}{=} \overline{\langle \alpha \mathbf{y}, \mathbf{x} \rangle} \overset{(S3)}{=} \overline{\alpha \langle \mathbf{y}, \mathbf{x} \rangle} = \overline{\alpha}\, \overline{\langle \mathbf{y}, \mathbf{x} \rangle} \overset{(S4)}{=} \overline{\alpha} \langle \mathbf{x}, \mathbf{y} \rangle$. ✓

Example 7.24. (a) Let $V = \mathbb{F}^n$.

Standard inner product $\mathbb{F}^n$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = x_1 \overline{y_1} + \ldots + x_n \overline{y_n} = \begin{pmatrix} \overline{y_1} & \cdots & \overline{y_n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad (7.13)$$

$$= \mathbf{y}^* \mathbf{x} =: \langle \mathbf{x}, \mathbf{y} \rangle_{\text{euclid}}, \qquad \mathbf{x}, \mathbf{y} \in \mathbb{F}^n$$

Again, the standard inner product is the most important one in $\mathbb{R}^n$ and $\mathbb{C}^n$. Since it describes the usual euclidean geometry, we denote it by $\langle \mathbf{x}, \mathbf{y} \rangle_{\text{euclid}}$ in both cases.

(b) For $V = \mathbb{F}^2$ and $\mathbf{x} = \binom{x_1}{x_2}$, $\mathbf{y} = \binom{y_1}{y_2} \in \mathbb{F}^2$ we define an inner product by

$$\langle \mathbf{x}, \mathbf{y} \rangle := \left\langle \binom{x_1}{x_2}, \binom{y_1}{y_2} \right\rangle := x_1 \overline{y_1} + x_1 \overline{y_2} + x_2 \overline{y_1} + 4 x_2 \overline{y_2}. = \left\langle \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} x, y \right\rangle_{euclid}$$ ✓

(c) For $V = \mathbb{F}^2$ and $\mathbf{x} = \binom{x_1}{x_2}$, $\mathbf{y} = \binom{y_1}{y_2} \in \mathbb{F}^2$, we could also define

$$\langle \mathbf{x}, \mathbf{y} \rangle = \left\langle \binom{x_1}{x_2}, \binom{y_1}{y_2} \right\rangle := x_1 \overline{y_2} + x_2 \overline{y_1}.$$

This is symmetric and linear in the first argument but not positive definite. For example, $\mathbf{x} = \binom{1}{-1}$ gives us $\langle \mathbf{x}, \mathbf{x} \rangle = \langle \binom{1}{-1}, \binom{1}{-1} \rangle = -2$.

(d) Let $V = \mathcal{P}([0,1], \mathbb{F})$ be the $\mathbb{F}$-vector space of all polynomial functions $\mathbf{f} : [0,1] \to \mathbb{F}$. Then, we define for $\mathbf{f}, \mathbf{g} \in V$ the inner product:

e.g. $f(x) = \cancel{x^2} + ix$ for $\mathbb{F} = \mathbb{C}$

$$\langle \mathbf{f}, \mathbf{g} \rangle := \int_0^1 \mathbf{f}(x) \overline{\mathbf{g}(x)}\, dx$$

$p(x) = ix$

$$\langle p, p \rangle = \int_0^1 p(x) \cdot \overline{p(x)}\, dx = \int_0^1 ix \cdot (-i) x\, dx = \int_0^1 x^2\, dx = \frac{1}{3} x^3 \Big|_0^1 = \frac{1}{3}$$

$\left(\sum x_i \cdot \overline{y_i} \rightsquigarrow \int \cdots dx\right.$

You should see the analogy to $\langle \mathbf{x}, \mathbf{y} \rangle_{\text{euclid}}$ in $\mathbb{F}^n$. All data is now continuously distributed over $[0,1]$, and we need an integral instead of a sum. Often, we are in the case $\mathbb{F} = \mathbb{R}$ and can ignore the complex conjugation $\overline{\mathbf{g}(x)}$.

Recall that for a general inner product on $\mathbb{R}^n$, there is a uniquely determined positive matrix A such that:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle A\mathbf{x}, \mathbf{y} \rangle_{\text{euclid}} \tag{7.14}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Symmetric if $\mathbb{F}=\mathbb{R}$

Definition 7.25. Positive definite matrix

A matrix $A \in \mathbb{F}^{n \times n}$ is called positive definite if it is selfadjoint ($A^* = A$) and satisfies

$$\langle A\mathbf{x}, \mathbf{x} \rangle_{\text{euclid}} > 0, \quad \textit{i.e.} \quad \mathbf{x}^* A\mathbf{x} > 0 \tag{7.15}$$

for all $\mathbf{x} \in \mathbb{F}^n \setminus \{\mathbf{o}\}$.

Attention! Positive definite needs selfadjointness

By our definition a positive definite matrix is always selfadjoint. In the complex case this follows from equation (7.15). However, in the real case, you cannot drop this assumption. Moreover, $\langle A\mathbf{x}, \mathbf{x} \rangle_{\text{euclid}}$ is always real, even in the case $\mathbb{F} = \mathbb{C}$,

$$\langle A\mathbf{x}, \mathbf{x} \rangle_{\text{euclid}} = \langle \mathbf{x}, A^*\mathbf{x} \rangle_{\text{euclid}} \overset{A=A^*}{=} \langle \mathbf{x}, A\mathbf{x} \rangle_{\text{euclid}} = \overline{\langle A\mathbf{x}, \mathbf{x} \rangle_{\text{euclid}}}.$$

Some authors might be using only equation (7.15) for defining positive definite matrices in the real case. Therefore to play it safe, we often talk about matrices that are "selfadjoint and positive definite".

Proposition 7.26. Positive definite matrix A $\Rightarrow$ $\langle A\mathbf{x}, \mathbf{y} \rangle_{\text{euclid}}$ inner product

If $A \in \mathbb{F}^{n \times n}$ is selfadjoint and positive definite, then

$$\langle \mathbf{x}, \mathbf{y} \rangle := \langle A\mathbf{x}, \mathbf{y} \rangle_{\text{euclid}}, \qquad \mathbf{x}, \mathbf{y} \in \mathbb{F}^n$$

defines an inner product in $\mathbb{F}^n$.

(Stretched geometry)

Example 7.27. Let us look at the examples from before:

(a) The identity matrix $\mathbb{1}$ is positive definite since $\langle \mathbb{1}\mathbf{x}, \mathbf{x}\rangle_{\text{euclid}} = \langle \mathbf{x}, \mathbf{x}\rangle_{\text{euclid}} > 0$ for all $\mathbf{x} \neq \mathbf{o}$.

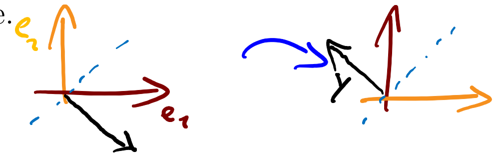
(b) The matrix $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 4 \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$ is positive definite since for all $\mathbf{x} = \left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) \in \mathbb{R}^2$ we have

$$= \left\langle \begin{pmatrix} x_1+x_2 \\ x_1+4x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\rangle_{euclid}$$

$$\left\langle \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}}_{A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\rangle_{\text{euclid}} = x_1x_1 + x_2x_1 + x_1x_2 + 4x_2x_2 = \underbrace{(x_1+x_2)^2}_{\geq 0} + 3\underbrace{(x_2)^2}_{\geq 0} \geq 0.$$

This can be only 0 if $x_1 = -x_2$ and $x_2 = 0$, hence only for $\mathbf{x} = \mathbf{o}$.

(c) The matrix $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ is selfadjoint but not positive definite. For example, for $\mathbf{x} = \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ the value $\langle A\mathbf{x}, \mathbf{x}\rangle_{\text{euclid}}$ is negative.



$\langle x, y\rangle_{eucl.} = \|x\| \cdot \|y\| \cdot \cos(\measuredangle(x,y)) < 0$ not positive definite

Proposition 7.28. 4 recognition features for a positive definite matrix

Let $A = (a_{ij}) \in \mathbb{F}^{n\times n}$ be a selfadjoint matrix. Then the following claims are equivalent:

i row $\to \begin{pmatrix} 1 & & \\ -\lambda & \ddots & \\ & & 1 \end{pmatrix}$

(i) *A is positive definite.*

$\Updownarrow$

(ii) *All eigenvalues A are positive.* > 0

(iii) *After using Gaussian elimination only with the matrices $Z_{i-\lambda j}$, all pivots are positive.* > 0

(iv) *The determinants of the so-called leading principal minors of A, which means $\det(H_1), \ldots, \det(H_n)$, are positive.* > 0

Here

$$H_1 = (a_{11}), \quad H_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad H_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad \ldots, \quad H_n = A.$$

Proof: Exercise! Use that selfadjoint matrices are unitarily diagonalizable.

Example 7.29. Let us check the proposition for the matrix $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 4 \end{smallmatrix}\right)$. It is positive definite by Example 7.27 (b). The eigenvalues of A are given by solving

$$0 = \det(A - \lambda\mathbb{1}) = (1-\lambda)(4-\lambda) - 1 = \lambda^2 - 5\lambda + 3, \quad \text{so} \quad \lambda_{1,2} = \underbrace{\frac{5}{2}}_{\text{“}2.5} \pm \underbrace{\sqrt{\left(\frac{5}{2}\right)^2 - 3}}. \quad > 0$$

$$\sqrt{\frac{25}{4} - \frac{12}{4}} = \sqrt{\frac{13}{4}} < \sqrt{\frac{16}{4}} = 2$$

Both eigenvalues, λ_1 and λ_2, are positive. The Gaussian elimination gives us:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} \boxed{1} & 1 \\ 0 & \boxed{3} \end{pmatrix}.$$

Both pivots, $\boxed{1}$ and $\boxed{3}$, are positive. At last the minors:

$$\det(H_1) = \det(1) = 1 > 0 \qquad \text{and} \qquad \det(H_2) = \det\begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} = 3 > 0.$$

Proposition 7.30. Inner products are related to pos. definite matrices

Let V be an $\mathbb{F}$-vector space with inner product $\langle\cdot,\cdot\rangle$ and $\dim(V) = n$. Let $\mathcal{B}$ be a basis of V. Then for all $\mathbf{x}, \mathbf{y} \in V$ we have

$\in \mathbb{F}^n$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle A\Phi_{\mathcal{B}}(\mathbf{x}), \Phi_{\mathcal{B}}(\mathbf{y}) \rangle_{\text{euclid}},$$

where $\langle\cdot,\cdot\rangle_{\text{euclid}}$ is the standard inner product in $\mathbb{F}^n$ and

$$A = G(\mathcal{B}) = \begin{pmatrix} \langle \mathbf{b}_1, \mathbf{b}_1 \rangle & \cdots & \langle \mathbf{b}_n, \mathbf{b}_1 \rangle \\ \vdots & & \vdots \\ \langle \mathbf{b}_1, \mathbf{b}_n \rangle & \cdots & \langle \mathbf{b}_n, \mathbf{b}_n \rangle \end{pmatrix}$$

is the Gramian matrix w.r.t. $\mathcal{B}$. (Cf. Homework 8.2)

Example 7.31. Look at the $\mathbb{R}$-vector space $\mathcal{P}_2([0,1])$ of all real polynomial functions $\mathbf{f} : [0,1] \to \mathbb{R}$ with degree ≤ 2. The integral

$$\langle \mathbf{p}, \mathbf{q} \rangle := \int_0^1 \mathbf{p}(x)\mathbf{q}(x)\, dx, \qquad \mathbf{p}, \mathbf{q} \in \mathcal{P}_2$$

$x^0 = 1$, x^1, x^2

defines an inner product. Let us check how to use Proposition 7.30 in this case. Choose a basis $\mathcal{B}$ of $\mathcal{P}_2$, for example the monomial basis $\mathcal{B} = (\mathbf{m}_0, \mathbf{m}_1, \mathbf{m}_2)$, and calculate the associated Gramian matrix:

$$\langle \mathbf{m}_i, \mathbf{m}_j \rangle = \int_0^1 x^i x^j\, dx = \int_0^1 x^{i+j}\, dx = \frac{x^{i+j+1}}{i+j+1}\Big|_0^1 = \frac{1^{i+j+1} - 0^{i+j+1}}{i+j+1} = \frac{1}{i+j+1} \tag{7.16}$$

and

$$G(\mathcal{B}) = \begin{pmatrix} \langle \mathbf{m}_0, \mathbf{m}_0 \rangle & \langle \mathbf{m}_1, \mathbf{m}_0 \rangle & \langle \mathbf{m}_2, \mathbf{m}_0 \rangle \\ \langle \mathbf{m}_0, \mathbf{m}_1 \rangle & \langle \mathbf{m}_1, \mathbf{m}_1 \rangle & \langle \mathbf{m}_2, \mathbf{m}_1 \rangle \\ \langle \mathbf{m}_0, \mathbf{m}_2 \rangle & \langle \mathbf{m}_1, \mathbf{m}_2 \rangle & \langle \mathbf{m}_2, \mathbf{m}_2 \rangle \end{pmatrix} \overset{(7.16)}{=} \begin{pmatrix} \frac{1}{0+0+1} & \frac{1}{1+0+1} & \frac{1}{2+0+1} \\ \frac{1}{0+1+1} & \frac{1}{1+1+1} & \frac{1}{2+1+1} \\ \frac{1}{0+2+1} & \frac{1}{1+2+1} & \frac{1}{2+2+1} \end{pmatrix} = \begin{pmatrix} 1/1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix}.$$

Then, by Proposition 7.30: For all $a, b, c, d, e, f \in \mathbb{R}$, we get:

$$\langle a\mathbf{m}_0 + b\mathbf{m}_1 + c\mathbf{m}_2 \,,\, d\mathbf{m}_0 + e\mathbf{m}_1 + f\mathbf{m}_2 \rangle = \left\langle \begin{pmatrix} 1/1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \begin{pmatrix} d \\ e \\ f \end{pmatrix} \right\rangle_{\text{euclid}}$$

$$= ad + \tfrac{1}{2}(ae + bd) + \tfrac{1}{3}(af + be + cd) + \tfrac{1}{4}(bf + ce) + \tfrac{1}{5}cf.$$

Let's check this:

$$\langle a\mathbf{m}_0 + b\mathbf{m}_1 + c\mathbf{m}_2 \,,\, d\mathbf{m}_0 + e\mathbf{m}_1 + f\mathbf{m}_2 \rangle = \int_0^1 (a + bx + cx^2)(d + ex + fx^2)\, dx$$

$$= \int_0^1 \Big(ad + (ae + bd)x + (af + be + cd)x^2 + (bf + ce)x^3 + cfx^4 \Big) dx$$

$$= ad \int_0^1 dx + (ae + bd) \int_0^1 x\, dx + (af + be + cd) \int_0^1 x^2 dx + (bf + ce) \int_0^1 x^3 dx + cf \int_0^1 x^4 dx$$

$$\overset{(7.16)}{=} ad + \tfrac{1}{2}(ae + bd) + \tfrac{1}{3}(af + be + cd) + \tfrac{1}{4}(bf + ce) + \tfrac{1}{5}cf.$$

Corollary 7.32. Gramian matrix is positive definite.

For a basis $\mathcal{B}$ of a vector space V with inner product $\langle \cdot, \cdot \rangle$, the Gramian matrix $G(\mathcal{B})$ is selfadjoint and positive definite.

Proof. $G(\mathcal{B}) = G(\mathcal{B})^*$ follows from $\langle \mathbf{b}_i, \mathbf{b}_j \rangle = \overline{\langle \mathbf{b}_j, \mathbf{b}_i \rangle}$. Using Proposition 7.30, we know $\langle G(\mathcal{B})\Phi_{\mathcal{B}}(\mathbf{x}), \Phi_{\mathcal{B}}(\mathbf{x}) \rangle_{\text{euclid}} = \langle \mathbf{x}, \mathbf{x} \rangle > 0$ for all $\mathbf{x} \in V \setminus \{\mathbf{o}\}$ and hence also for all vectors $\Phi_{\mathcal{B}}(\mathbf{x}) \in \mathbb{F}^n \setminus \{\mathbf{o}\}$. □

7.5.2 Norms

As always, let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and V be an $\mathbb{F}$-vector space. Even in the case V not having an inner product, we can talk about the length of vectors if we define a length measure:

Definition 7.33. Norm

A map $\|\cdot\| : V \to \mathbb{R}$ with the following properties is called a norm on V. For all $\mathbf{x}, \mathbf{y} \in V$ and $\alpha \in \mathbb{F}$, we have:

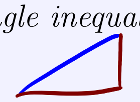
(N1) $\|\mathbf{x}\| \geq 0$, *and* $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{o}$, *(positive definite)*

(N2) $\|\alpha\mathbf{x}\| = |\alpha| \, \|\mathbf{x}\|$, *(absolutely homogeneous)*

abs. value in $\mathbb{R}$ or $\mathbb{C}$

(N3) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ *(triangle inequality).*

An $\mathbb{F}$-vector space with such a norm is called a normed space.



Example 7.34. (a) We already know that the euclidean norm for $\mathbb{F}^n$, given by

$$\|\mathbf{x}\| = \left\| \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\| = \sqrt{|x_1|^2 + \cdots + |x_n|^2}, \qquad \mathbf{x} \in \mathbb{F}^n, \tag{7.17}$$

absolute value in $\mathbb{C}$ (or $\mathbb{R}$)

satifies (N1-3) from Definition 7.33.

(b) In equation (7.17), you see squares and a square root that cancel themselves in some sense. This would also work for cubes and the third root. Or even in general:

The p-norm

For each real number $p \geq 1$, we set:

$$\|\mathbf{x}\|_p = \left\| \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\|_p := \sqrt[p]{|x_1|^p + \cdots + |x_n|^p}, \qquad \mathbf{x} \in \mathbb{F}^n. \tag{7.18}$$

This defines the so-called *p-norm*. The euclidean norm (7.17) is hence also called 2-norm.

(c) Another related norm is given by:

$$\lim_{p \to \infty} \sqrt[p]{|x_1|^p + \cdots + |x_n|^p} = \max\{|x_1|, \ldots, |x_n|\}$$

analysis next semester

Maximum norm or ∞-norm

$$\|\mathbf{x}\|_\infty = \left\| \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\|_\infty := \max\{|x_1|, \ldots, |x_n|\} \tag{7.19}$$

$$= \lim_{p\to\infty} \|\mathbf{x}\|_p, \qquad \mathbf{x} \in \mathbb{F}^n.$$

$\|x\|_{max}$ =

Let us check for $n = 2$ that the three properties in Definition 7.33. Let $\alpha \in \mathbb{F}$ and $\mathbf{x} = \binom{x_1}{x_2}, \mathbf{y} = \binom{y_1}{y_2} \in \mathcal{F}^2$.

(N1) $\|\mathbf{x}\|_\infty = \max\{|x_1|, |x_2|\}$ is only 0 if $x_1 = 0$ and $x_2 = 0$, hence $\mathbf{x} = \mathbf{o}$.

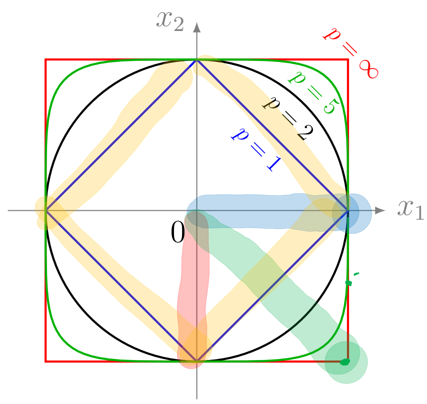
(N2) $\|\alpha\mathbf{x}\|_\infty = \max\{|\alpha x_1|, |\alpha x_2|\} = \max\{|\alpha|\,|x_1|, |\alpha|\,|x_2|\} = |\alpha| \max\{|x_1|, |x_2|\} = |\alpha|\|\mathbf{x}\|_\infty$

(N3) The triangle inequality:

$$\|\mathbf{x}+\mathbf{y}\|_\infty = \max\{|x_1+y_1|, |x_2+y_2|\} \leq \max\{|x_1|+|y_1|, |x_2|+|y_2|\}$$

$$\overset{(*)}{\leq} \max\{|x_1|, |x_2|\} + \max\{|y_1|, |y_2|\} = \|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty$$

(*) use largest number on both sides

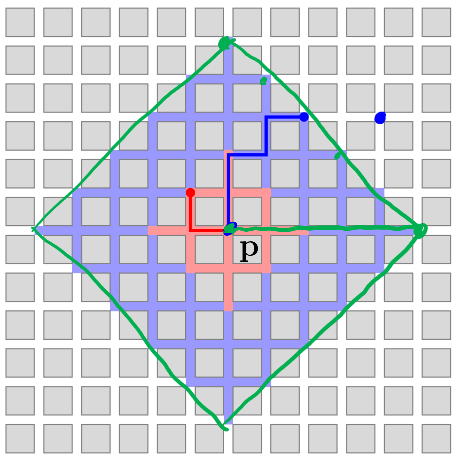


On the right-hand side, you see the geometric picture for different norms. Usually, one calls it the "unit circles", which means the sets

= unit balls

$$\{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\|_p = 1\}$$

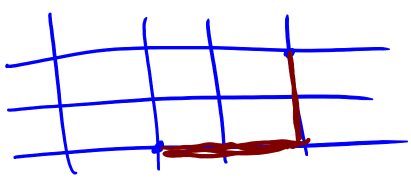
Such a subset of $\mathbb{R}^2$ consists of all vectors with length 1, for different $p = 1$, 2, 5 and ∞.
For $p = 2$, this in indeed a usual circle. However, also the different geometric views for other p are interesting:



Assume you are in Manhattan inside a taxicab at point $\mathbf{p}$. Driving one block costs you \$ 1. If you have \$ 2 in your pocket, you can reach all the red points in the map. If you have \$ 5, you can get to all the red and the blue points. The ε-neighbourhoods

$$\{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}-\mathbf{p}\| < \varepsilon\}$$

in Manhattan are just squares, which stay at one corner, and not real circles. This exactly the 1-norm $\|\cdot\|_1$, which is often alternatively called "taxicab norm".



Rule of thumb: Norm gives you lengths and distances

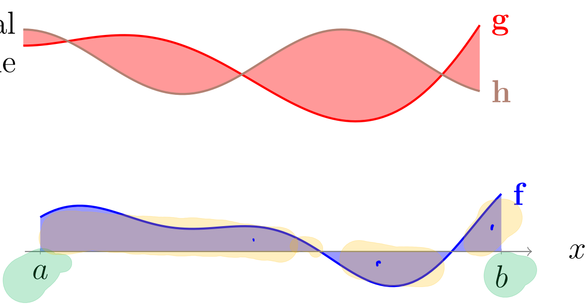
You should imagine $\|\mathbf{x}\|$ as the length of the "vector arrow" $\mathbf{x}$. Hence, $\|\mathbf{x}-\mathbf{y}\|$ is the length of the connection vector between $\mathbf{x}$ and $\mathbf{y}$ – or in other words: The distance between $\mathbf{x}$ and $\mathbf{y}$.

Example 7.35. – p-norm for polynomials. The p-norms in $\mathbb{F}^n$, which we defined above, can be generalised for functions. For example, for the $\mathbb{R}$-vector space $\mathcal{P}([a,b])$, which means all polynomial functions $\mathbf{f} : [a,b] \to \mathbb{R}$, we can also define such norms:

Norms for polynomials on $[a,b]$

$$\|\mathbf{f}\|_p := \sqrt[p]{\int_a^b |\mathbf{f}(x)|^p \, dx} \quad \textit{for } p \in [1,\infty) \qquad \textit{and} \qquad \|\mathbf{f}\|_\infty := \max_{x\in[a,b]} |\mathbf{f}(x)|$$

On the right-hand side, you see some polynomial functions $\mathbf{f},\mathbf{g},\mathbf{h} \in \mathcal{P}([a,b])$. The area of the blue region is



$$\|\mathbf{f}\|_1 = \int_a^b |\mathbf{f}(x)| \, dx$$

and the area of the red region is:

$$\|\mathbf{g}-\mathbf{h}\|_1 = \int_a^b |\mathbf{g}(x) - \mathbf{h}(x)| \, dx.$$

In later lectures, like mathematical analysis, we will prove the three properties (N1),(N2) and (N3) for all these norms.

7.5.3 Norm in pre-Hilbert spaces

Proposition & Definition 7.36. Induced or associated norm

Let V be a pre-Hilbert space, which is an $\mathbb{F}$-vector space with an inner product $\langle\cdot,\cdot\rangle$. Then

$$\|\mathbf{x}\| := \sqrt{\langle\mathbf{x},\mathbf{x}\rangle}, \qquad \mathbf{x} \in V$$

defines a norm and it is called the induced norm or associated norm w.r.t. $\langle\cdot,\cdot\rangle$.

Proposition 7.37. Cauchy-Schwarz inequality

Let V be a pre-Hilbert space. For all $\mathbf{x}, \mathbf{y} \in V$:

$$|\langle \mathbf{x}, \mathbf{y} \rangle|^2 \leq \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle.$$

With the associated norm from Proposition& Definition 7.36, we get:

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|.$$

Equality holds if and only if $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent.

Example 7.38. (a) The standard inner product $\langle \mathbf{x}, \mathbf{y} \rangle_{\text{euclid}} = x_1 \overline{y_1} + \cdots + x_n \overline{y_n}$ in $\mathbb{F}^n$ induced the 2-norm $\|\mathbf{x}\| = \sqrt{|x_1|^2 + \cdots + |x_n|^2}$ in $\mathbb{F}^n$.

(b) The associated norm with respect to the inner product $\langle \mathbf{x}, \mathbf{y} \rangle := \langle A\mathbf{x}, \mathbf{y} \rangle_{\text{euclid}}$ in $\mathbb{F}^n$ where $A \in \mathbb{F}^{n \times n}$ is a selfadjoint and positive definite matrix is given by

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\langle A\mathbf{x}, \mathbf{x} \rangle_{\text{euclid}}}.$$

For the example $A = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$, we get

$$\left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\| = \sqrt{\left\langle \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\rangle_{\text{euclid}}} = \sqrt{|x_1|^2 + x_1 \overline{x_2} + x_2 \overline{x_1} + 4|x_2|^2}.$$

(c) Looking at the $\mathbb{F}$-vector space $\mathcal{P}([a,b])$, $\mathbb{F}$) of all polynomial functions $\mathbf{f} : [a,b] \to \mathbb{F}$, we defined the inner product

$$\langle \mathbf{f}, \mathbf{g} \rangle = \int_a^b \mathbf{f}(x) \overline{\mathbf{g}(x)} \, dx \,. \tag{7.20}$$

The associated norm in $\mathcal{P}([a,b])$ is the already introduced 2-norm since

$$\|\mathbf{f}\| = \sqrt{\langle \mathbf{f}, \mathbf{f} \rangle} = \sqrt{\int_a^b \mathbf{f}(x) \overline{\mathbf{f}(x)} \, dx} = \sqrt{\int_a^b |\mathbf{f}(x)|^2 \, dx} = \|\mathbf{f}\|_2.$$

7.5.4 Recollection: Angles, orthogonality and projection

Let V be a pre-Hilbert space, which means an $\mathbb{F}$-vector space with given inner product $\langle \cdot, \cdot \rangle$, and let $\|.\|$ be the associated norm.

In this case, we have again the geometric structure and can talk about angles, orthogonal vectors and orthogonal projections:

Proposition & Definition 7.39. Still the same about orthogonality:

- *For* $\mathbf{x}, \mathbf{y} \in V$ *we write* $\mathbf{x} \perp \mathbf{y}$ *if* $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.
- *For* $\mathbb{F} = \mathbb{R}$ *and* $\mathbf{x}, \mathbf{y} \in V \setminus \{\mathbf{o}\}$ *we define:*

$$\operatorname{angle}(\mathbf{x}, \mathbf{y}) := \arccos\left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}\right).$$

- *For a nonempty set* $M \subset V$ *we call*

$$M^\perp := \{\mathbf{x} \in V : \mathbf{x} \perp \mathbf{m} \text{ for all } \mathbf{m} \in M\}$$

the orthogonal complement of M. *This is always a subspace of* V. *Instead of* $\mathbf{x} \in M^\perp$, *we often write* $x \perp M$.

- *For* $\mathbf{x} \in V$ *and a subspace* U *of* V *there is a unique decomposition*

$$\mathbf{x} = \mathbf{p} + \mathbf{n} =: \mathbf{x}_{|U} + \mathbf{x}_{|U^\perp}$$

into the orthogonal projection $\mathbf{p} =: \mathbf{x}_{|U} \in U$ *and the normal component* $\mathbf{n} = \mathbf{x}_{|U^\perp} \in U^\perp$ *with respect to* U. *The calculation is given by*

$$G(\mathcal{B})\, \Phi_{\mathcal{B}}(\mathbf{p}) = \begin{pmatrix} \langle \mathbf{x}, \mathbf{b}_1 \rangle \\ \vdots \\ \langle \mathbf{x}, \mathbf{b}_n \rangle \end{pmatrix} \qquad (7.21)$$

$$\Rightarrow p = \Phi_{\mathcal{B}}^{-1} G(\mathcal{B})^{-1} \begin{pmatrix} \langle x, b_1 \rangle \\ \vdots \\ \langle x, b_n \rangle \end{pmatrix}$$

for any basis $\mathcal{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ *of* U, *and* $\mathbf{n} = \mathbf{x} - \mathbf{p}$.

- *A family* $\mathcal{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ *with vectors from* V *is called:*
  - *Orthogonal system (OS) if* $\mathbf{u}_i \perp \mathbf{u}_j$ *for all* $i, j = 1, \ldots, n$ *with* $i \neq j$;
  - *Orthonormal system (ONS) if, in addition,* $\|\mathbf{u}_i\| = 1$ *for all* $i = 1, \ldots, n$;
  - *Orthogonal basis (OB) if it an OS and a basis of* V;
  - *Orthonormal basis (ONB) if it an ONS and a basis of* V.
- *OS that do not own the zero vector* $\mathbf{o}$ *are always linearly independent.*
- *If* $\mathcal{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ *is an OB of* U, *then the equation (7.21) is much simpler:*

$$\Phi_{\mathcal{B}}(\mathbf{x}_{|U}) = \begin{pmatrix} \frac{\langle \mathbf{x}, \mathbf{b}_1 \rangle}{\|\mathbf{b}_1\|^2} \\ \vdots \\ \frac{\langle \mathbf{x}, \mathbf{b}_n \rangle}{\|\mathbf{b}_n\|^2} \end{pmatrix}, \quad i.e. \quad \mathbf{x}_{|U} = \frac{\langle \mathbf{x}, \mathbf{b}_1 \rangle}{\|\mathbf{b}_1\|^2} \mathbf{b}_1 + \ldots + \frac{\langle \mathbf{x}, \mathbf{b}_n \rangle}{\|\mathbf{b}_n\|^2} \mathbf{b}_n. \qquad (7.22)$$

If $\mathcal{B}$ *is an ONB, then it gets also easier* $\|\mathbf{b}_i\|^2$ $(= 1)$.

Remember → $x_{|U} = \langle x, b_1 \rangle b_1 + \langle x, b_2 \rangle b_2 + \cdots + \langle x, b_n \rangle b_n$

Example 7.40. (a) The vectors $\mathbf{x} = \binom{1}{\mathrm{i}}$ and $\mathbf{y} = \binom{0}{1}$ from $\mathbb{C}^2$ are not orthogonal w.r.t. the standard inner product $\langle \cdot, \cdot \rangle_{\text{euclid}}$ since

$$\left\langle \begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle_{\text{euclid}} = 1 \cdot \overline{0} + \mathrm{i} \cdot \overline{1} = \mathrm{i} \neq 0.$$