Chapter 5

Norms & Procrustes problems

Contents (class version)

5.0 Introduction

This chapter discusses **vector norms** and **matrix norms** (also known as **operator norms**) in more generality and applies them to the **Procrustes problem**. Source material for this chapter includes [1, §7.2-7.4].

So far the norms we have considered are the **Euclidean norm** $\|\boldsymbol{x}\|_2$, the **spectral norm** $\|\boldsymbol{A}\|_2$, and the **Frobenius norm** $\|\boldsymbol{A}\|_{\mathrm{F}}$. These three norms are particularly important, but there are many other important norms for signal processing applications.

The Tikhonov regularized least-squares problem in Ch. 4 illustrates the two primary uses of norms:

$$\hat{x} = \arg\min_{x \in \mathbb{F}^N} \underbrace{\|Ax - y\|_2^2}_{\substack{\text{distance} \\ \text{“how far”}}} + \beta \underbrace{\|x\|_2^2}_{\substack{\text{size} \\ \text{“how big”}}}$$

5.1 Vector norms

L§7.3

So far the only **vector norm** discussed in these notes has been the common **Euclidean norm**.
Many other vector norms are also important in signal processing and machine learning.

Define. [1, p. 57] A norm on a vector space $\mathcal{V}$ defined over a field $\mathbb{F}$ is a function $\|\cdot\|$ from $\mathcal{V}$ to $[0, \infty)$ that satisfies the following properties $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$:

- $\|x\| \geq 0$ (nonnegative) ✓
- $\|x\| = \underset{\mathbb{R}}{0} \iff x = \underline{0} \in \mathcal{V}$ (positive) ✓
- $\|\alpha x\| = |\alpha| \, \|x\|$ (homogeneous)
- $\|x + y\| \leq \|x\| + \|y\|$ (**triangle inequality**)

Examples of vector norms

- For $1 \leq p < \infty$, the ℓ_p **norm** is

$$\|\boldsymbol{x}\|_p \triangleq \left(\sum_i |x_i|^p\right)^{1/p} \tag{5.1}$$

  - The **vector 2-norm** or **Euclidean norm** is the case $p = 2$: $\|\boldsymbol{x}\|_2 \triangleq \sqrt{\sum_i |x_i|^2}$.
  - The 1-norm or "Manhattan norm" is the case $p = 1$: $\|\boldsymbol{x}\|_1 \triangleq \sum_i |x_i|$.

- The **max norm** or **infinity norm** or ℓ_∞ **norm** is

$$\|\boldsymbol{x}\|_\infty \triangleq \sup\{|x_1|, |x_2|, \ldots\}, \tag{5.2}$$

where sup denotes the **supremum** (least upper bound) of a set. One can show [2, Prob. 2.12] that

$$\|\boldsymbol{x}\|_\infty = \lim_{p\to\infty} \|\boldsymbol{x}\|_p. \tag{5.3}$$

For the vector space $\mathbb{F}^N$, the supremum is simply a maximum:

$$\boxed{\|\boldsymbol{x}\|_\infty \triangleq \max\{|x_1|, \ldots, |x_N|\}.} \tag{5.4}$$

- For quantifying (non)sparsity of a vector, it is useful to note that

$$\lim_{p\to 0} \|\boldsymbol{x}\|_p^p = \sum_i \mathbb{1}_{\{x_i \neq 0\}} = \|x\|_0 \tag{5.5}$$

where $\mathbb{I}_{\{\cdot\}}$ denotes the **indicator function** that is unity if the argument is true and zero if false. However, the "0-norm" $\|\boldsymbol{x}\|_0$ is *not* a vector norm because it does not satisfy the at least one of the conditions of the norm definition above. The proper name for $\|\boldsymbol{x}\|_0$ is **counting measure**.

- Sometimes we want a **weighted** norm, *e.g.*, the **weighted 2-norm** is

$$\|\boldsymbol{x}\|_{\boldsymbol{W}} = \sqrt{(\boldsymbol{x}'\boldsymbol{W}\boldsymbol{x})}.$$

Exercise. Show that the weighted Euclidean norm $\|\boldsymbol{x}\|_{\boldsymbol{W}}$ is a norm iff $\boldsymbol{W}$ is a positive definite matrix. In particular, if $\boldsymbol{W}$ is a $N \times N$ diagonal matrix with positive diagonal elements w_i, then

$$\|\boldsymbol{x}\|_{\boldsymbol{W}} = \left(\sum_{i=1}^N w_i \, |x_i|^2\right)^{1/2}. = \left(\sum_i |\sqrt{w_i}\, x_i|^2\right)^{1/2}$$

1. Which of the four properties of a vector norm does the counting measure $\|\cdot\|_0$ satisfy?

A: 1,2 B: 1,3 C: 1,2,3 D: 1,2,4 E: 1,3,4 ??

$$\|x+y\|_0 = \sum_i \mathbb{1}_{\{x_i+y_i \neq 0\}} \le \sum_i \mathbb{1}_{\{x_i \neq 0\}} + \sum_i \mathbb{1}_{\{y_i \neq 0\}} = \|x\|_0 + \|y\|_0$$

Practical implementation

For the preceding examples, in JULIA, first invoke `using LinearAlgebra`
then use:

$\|v\|_p$ `norm(v, p)`
$\|v\|_2$ `norm(v, 2)` or just `norm(v)`
$\|v\|_1$ `norm(v, 1)`
$\|v\|_\infty$ `norm(v, Inf)`
$\|v\|_0$ `norm(v, 0)` = count(!= (0), v)

Caution. For $p < 1$, $\|\cdot\|_p$ is not a proper vector **norm**, though it is sometimes used in practical problems and `norm(v, p)` will evaluate (5.1) for any $-\infty \le p \le \infty$.

2. If W is `Diagonal(w)` for $w > 0$, then which command computes the weighted norm $\|x\|_W$?

A: `norm(x .* w)`
B: `norm(x .* w, 2)`
C: `norm(x .* w.^2)`
D: `norm(x .* sqrt.(w))`
E: None

??

2021-10-07 //

Properties of norms

- Let $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ be any two vector norms on a *finite-dimensional* space. Then there exist finite positive constants C_m and C_M (that depend on α and β) such that:

$$C_m \|\cdot\|_\alpha \leq \|\cdot\|_\beta \leq C_M \|\cdot\|_\alpha . \tag{5.6}$$

In a sense then "all norms are equivalent" to within constant factors. This equivalence is especially relevant when examining convergence of a sequence (see p. 5.42); if a sequence converges w.r.t. one norm, then it converges w.r.t. any norm.

- For any vector norm, the **reverse triangle inequality** is:

$$\begin{matrix} \|x\| - \|y\| \\ \|y\| - \|x\| \end{matrix} \leq \left| \|\boldsymbol{x}\| - \|\boldsymbol{y}\| \right| \leq \|\boldsymbol{x} - \boldsymbol{y}\|, \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}.$$

Proof: $\|\boldsymbol{x}\| = \|\boldsymbol{x} - \boldsymbol{y} + \boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{y}\| + \|\boldsymbol{y}\| \Longrightarrow \|\boldsymbol{x}\| - \|\boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$.
Similarly $\|\boldsymbol{y}\| - \|\boldsymbol{x}\| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$. Now combine these two inequalities.

- Any vector norm $\|\cdot\|$ on a vector space $\mathcal{V}$ is a **convex function**:

$$\|\alpha x + (1-\alpha) y\| \leq \alpha \|x\| + (1-\alpha) \|y\| \qquad \forall \alpha \in [0,1] \quad \forall x, y \in \mathcal{V}$$

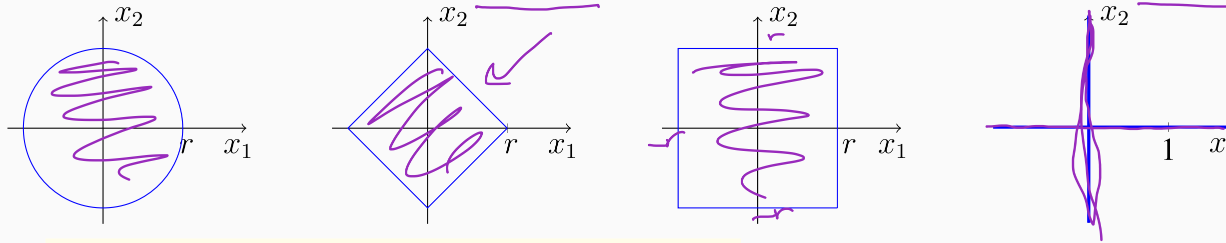
This fact is easy to prove using the **triangle inequality** and the homogeneity property. (HW)

- For $p > 1$, the function $f(\boldsymbol{x}) \triangleq \|\boldsymbol{x}\|_p^p$ is **strictly convex**.

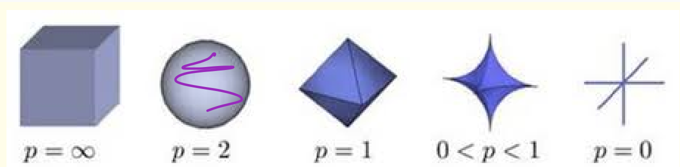
- For any norm, the ball of radius $r > 0$ $\{\boldsymbol{x} : \|\boldsymbol{x}\| \leq r\}$ is a **convex set**. (HW)

Example. Here is an illustration of some norm balls in $\mathbb{R}^2$.

$\{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x}\|_2 \leq r\}$ $\{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x}\|_1 \leq r\}$ $\{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x}\|_\infty \leq r\}$ $\{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x}\|_0 \leq 1\}$



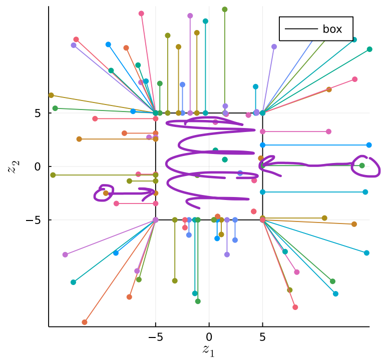
In $\mathbb{R}^3$:



3. For $\mathcal{C} = \left\{ \boldsymbol{x} \in \mathbb{R}^N \ : \ \|\boldsymbol{x}\|_\infty \leq 5 \right\}$, which of these is the **projection** of a point $\boldsymbol{z} \in \mathbb{R}^N$ onto $\mathcal{C}$?

A: `min.(z,5)` B: `min.(abs.(z),5)` C: `min.(abs.(z),5).*sign.(z)` D: None of these

??



$\mathbb{R}^2$

challenge?

$\mathbb{C}^N$?

"box constraint"

Norm notation

Some math literature uses $|\boldsymbol{x}|$ instead of $\|\boldsymbol{x}\|$ to denote a vector norm.
That notation should be avoided for matrices where $|\boldsymbol{A}|$ often denotes the determinant of $\boldsymbol{A}$.
Sometimes one must determine from context what $|\cdot|$ means in such literature.

Unitarily invariant vector norms

Some vector norms have the following useful property.

Define. A vector norm $\|\cdot\|$ on $\mathbb{F}^N$ is **unitarily invariant** iff for every unitary matrix $\boldsymbol{U} \in \mathbb{F}^{N \times N}$:

$$\|Ux\| = \|x\| \quad \forall x \in \mathbb{F}^N$$

Example. The Euclidean norm $\|\cdot\|_2$ on $\mathbb{F}^N$ is unitarily invariant, because for any unitary $\boldsymbol{U}$ (see p. 1.68):

$$\|\boldsymbol{U}\boldsymbol{x}\|_2 = \sqrt{(\boldsymbol{U}\boldsymbol{x})'(\boldsymbol{U}\boldsymbol{x})} = \sqrt{\boldsymbol{x}'\boldsymbol{U}'\boldsymbol{U}\boldsymbol{x}} = \sqrt{\boldsymbol{x}'\boldsymbol{x}} = \|\boldsymbol{x}\|_2, \ \forall \boldsymbol{x} \in \mathbb{F}^N.$$

As noted previously, this property is related to **Parseval's theorem**.

Example. $\|\cdot\|_1$ is not unitarily invariant.
If $\boldsymbol{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ and $\boldsymbol{x} = \boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ then $\|\boldsymbol{x}\|_1 = 1$ but $\|\boldsymbol{U}\boldsymbol{x}\|_1 = \sqrt{2}$.

Exercise. Examine the "Fourier sparsity loss" in [3] and consider the implications of **unitary invariance**.

Example. Another unitary invariant norm on $\mathbb{F}^N$ is $\|\cdot\|_{(\alpha)} \triangleq \alpha \|\cdot\|_2$ for any $\alpha > 0$.

Challenge. Find another unitarily invariant norm on $\mathbb{F}^N$ or prove that no others exist. ♦♦

Inner products

L§7.2

Most of the vector spaces used in this course are **inner product spaces**, meaning a vector space with an associated **inner product** operation.

Define. For a vector space $\mathcal{V}$ over the field $\mathbb{F}$, an **inner product** operation is a function $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{F}$ that must satisfy the following axioms $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}, \ \alpha \in \mathbb{F}$.

$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle^*$ (**Hermitian symmetry**)

$\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$ (**additivity**)

$\langle \alpha \boldsymbol{x}, \boldsymbol{y} \rangle = \alpha \langle \boldsymbol{x}, \boldsymbol{y} \rangle$ (**scaling**)

$\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$ and $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ iff $\boldsymbol{x} = \boldsymbol{0}$. (**positive definite**)

If $A \succ 0$

then

valid inner product!

$\langle x, y \rangle = y' A x$

$\langle x, x \rangle = x' A x > 0 \quad \forall x \neq 0$

Examples of inner products

Example. For vectors in $\mathbb{F}^N$, the usual inner product is

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{n=1}^{N} x_n y_n^*.$$

Example. For the (infinite dimensional) vector space of square integrable functions on the interval $[a, b]$, the following integral is a valid inner product:

cf Fourier series!

$$\langle f, g \rangle = \int_a^b f(t) g^*(t) \, \mathrm{d}t \, .$$

♦♦

An inner product for random variables (Read)

Example. For two real, zero-mean random variables X, Y defined on a joint probability space, a natural inner product is $\mathsf{E}[XY]$. (Keep in mind that random variables are functions.) With this definition, the corresponding norm is $\|X\| \triangleq \sqrt{\langle X, X \rangle} = \sqrt{\mathsf{E}[X^2]} = \sigma_X$, the standard deviation of X. Here, the **Cauchy-Schwarz inequality** is equivalent to usual bound on the **correlation coefficient**: $\rho_{X,Y} \triangleq \frac{\mathsf{E}[XY]}{\sigma_X \sigma_Y} \Longrightarrow |\rho_{X,Y}| \leq 1$. ♦♦

Exercise. With this definition of inner product, what types of random variables are "orthogonal"? Pairs of random variables that are **uncorrelated**, *i.e.*, where $\mathsf{E}[XY] = 0$.

Example. For two matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{F}^{M \times N}$ (a **vector space**!), the **Frobenius inner product**, also called the **Hilbert–Schmidt inner product**, is defined as:

$$\langle \boldsymbol{A}, \boldsymbol{B} \rangle \triangleq \operatorname{trace}(A B') = \sum_{m=1}^{M} [A B']_{mm} = \sum_{m=1}^{M} \sum_{n=1}^{N} a_{mn} b_{mn}^* \tag{5.7}$$

$\langle \operatorname{vec}(A), \operatorname{vec}(B) \rangle$

where the latter inner product is the ordinary Euclidean inner product on $\mathbb{F}^{MN}$.

Exercise. Verify the four properties above for these inner product examples.

The **induced norm** corresponding to the **Frobenius inner product** is the **Frobenius norm**:

$$\|\boldsymbol{A}\|_{\mathrm{F}} = \sqrt{\langle \boldsymbol{A},\, \boldsymbol{A} \rangle}. \quad = \sqrt{\langle \mathrm{vec}(A), \mathrm{vec}(A) \rangle}$$

Properties of inner products

- Bilinearity:

$$\left\langle \sum_i \alpha_i \boldsymbol{x}_i,\, \sum_j \beta_j \boldsymbol{y}_j \right\rangle = \sum_i \sum_j \alpha_i \beta_j^* \langle \boldsymbol{x}_i,\, \boldsymbol{y}_j \rangle, \qquad \forall \{\boldsymbol{x}_i\}, \{\boldsymbol{y}_j\} \in \mathcal{V}.$$

$\forall \{\alpha_i\}, \{\beta_j\} \in \mathbb{F}$

- Any valid vector **inner product** induces a valid vector **norm**:

$$\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x},\, \boldsymbol{x} \rangle}. \tag{5.8}$$

Exercise. Verify that such an **induced norm** satifies the four conditions for a norm on p. 5.3.

- A vector norm satisfies the **parallelogram law**: ♦♦

$$\frac{1}{2}\left(\|\boldsymbol{x}+\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2\right) = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2, \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V},$$

iff it is induced by an inner product via (5.8). The required inner product is

$$\begin{aligned}\langle \boldsymbol{x},\, \boldsymbol{y} \rangle &\triangleq \frac{1}{4}\left(\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2 + \imath \|\boldsymbol{x}+\imath\boldsymbol{y}\|^2 - \imath \|\boldsymbol{x}-i\boldsymbol{y}\|^2\right)\\ &= \frac{\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2}{2} + \imath \frac{\|\boldsymbol{x}+\imath\boldsymbol{y}\|^2 - \|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2}{2}.\end{aligned}$$

- The **Cauchy-Schwarz inequality** (or **Schwarz** or **Cauchy-Bunyakovsky-Schwarz** inequality) states:

$$|\langle x, y \rangle| \le \|x\| \, \|y\| \quad \forall x, y \in \mathcal{V} \tag{5.9}$$

for a norm $\|\cdot\|$ induced by an inner product $\langle \cdot, \cdot \rangle$ via (5.8), with equality iff $\boldsymbol{x}$ and $\boldsymbol{y}$ are linearly dependent.

Example. Applying the inequality to the **Frobenius inner product** for matrices yields:

$$|\mathrm{trace}\{\boldsymbol{A}\boldsymbol{B}'\}| = |\langle \boldsymbol{A}, \boldsymbol{B} \rangle| \le \|\boldsymbol{A}\|_{\mathrm{F}} \, \|\boldsymbol{B}\|_{\mathrm{F}} \,.$$

4. In an inner product space on $\mathbb{R}^N$, is $\langle \boldsymbol{x}, \boldsymbol{y} \rangle \le \|\boldsymbol{x}\| \, \|\boldsymbol{y}\|$? where $\|\cdot\|$ is the induced norm

A: Yes, always B: Not always C: Never ??

Proof of Cauchy-Schwarz inequality for $\mathbb{F}^N$ $t \le |t| \quad \forall t \in \mathbb{R}$ (Read)

For any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^N$ let $\boldsymbol{A} = [\boldsymbol{x} \; \boldsymbol{y}]$, so $\boldsymbol{A}'\boldsymbol{A} = \begin{bmatrix} \boldsymbol{x}'\boldsymbol{x} & \boldsymbol{x}'\boldsymbol{y} \\ \boldsymbol{y}'\boldsymbol{x} & \boldsymbol{y}'\boldsymbol{y} \end{bmatrix}$.

$\boldsymbol{A}'\boldsymbol{A}$ is Hermitian symmetric $\Longrightarrow$ its eigenvalues are all real and nonnegative.
$\Longrightarrow \det\{\boldsymbol{A}'\boldsymbol{A}\} \ge 0 \Longrightarrow (\boldsymbol{x}'\boldsymbol{x})(\boldsymbol{y}'\boldsymbol{y}) - (\boldsymbol{y}'\boldsymbol{x})(\boldsymbol{x}'\boldsymbol{y}) \ge 0 \Longrightarrow |\boldsymbol{x}'\boldsymbol{y}|^2 \le (\boldsymbol{x}'\boldsymbol{x})(\boldsymbol{y}'\boldsymbol{y}) = \|\boldsymbol{x}\|_2^2 \, \|\boldsymbol{y}\|_2^2 \,.$
Taking the square root of both sides yields the inequality. □

We used the fact that $(\boldsymbol{y}'\boldsymbol{x})(\boldsymbol{x}'\boldsymbol{y}) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle \langle \boldsymbol{y}, \boldsymbol{x} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle (\langle \boldsymbol{x}, \boldsymbol{y} \rangle)^* = |\langle \boldsymbol{x}, \boldsymbol{y} \rangle|^2 \,.$

Angle between vectors (Read)

Define. The **angle** θ between two nonzero vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$ w.r.t. inner product $\langle \cdot, \cdot \rangle$ having induced norm $\|\cdot\|$ is defined by

$$\cos\theta = \frac{|\langle x, y\rangle|}{\|x\| \, \|y\|} \Rightarrow \theta \in [0, \pi/2]$$

For real vectors, we can omit the absolute value and obtain $\theta \in [0, \pi]$.

The **Cauchy-Schwarz inequality** is equivalent to the statement $|\cos\theta| \leq 1$.

Ch. 1 introduced this inequality in Euclidean space, but in fact it holds in any inner product space.

Exercise. Determine the angle in $\mathbb{F}^{N \times N}$ between $\boldsymbol{I}_N$ and $\boldsymbol{x}\boldsymbol{x}'$ for $\boldsymbol{x} \in \mathbb{F}^N - \{\mathbf{0}\}$. $\cos\theta = |\langle \boldsymbol{I}, \boldsymbol{x}\boldsymbol{x}' \rangle| / (\|\boldsymbol{I}\|_{\mathrm{F}} \, \|\boldsymbol{x}\boldsymbol{x}'\|_{\mathrm{F}}) = \mathsf{trace}\{\boldsymbol{x}\boldsymbol{x}'\} / (N \, \|\boldsymbol{x}\|_2^2) = 1/N$

Angle between subspaces (Read)

Define. The angle between two subspaces $\mathcal{S}$ and $\mathcal{T}$ of a vector space $\mathcal{V}$ is the minimum angle between nonzero vectors in those subspaces [4] and has the following cosine:

$$\cos\theta = \underset{\boldsymbol{s}\in\mathcal{S}-\{\boldsymbol{0}\},\, \boldsymbol{t}\in\mathcal{T}-\{\boldsymbol{0}\}}{\overset{\max}{\cancel{\min}}} \frac{|\langle \boldsymbol{s},\, \boldsymbol{t}\rangle|}{\|\boldsymbol{s}\|_2 \, \|\boldsymbol{t}\|_2} = \underset{\boldsymbol{s}\in\mathcal{S},\, \boldsymbol{t}\in\mathcal{T}}{\overset{\max}{\cancel{\min}}} |\langle \boldsymbol{s},\, \boldsymbol{t}\rangle| \text{ s.t. } \|\boldsymbol{s}\|_2 = \|\boldsymbol{t}\|_2 = 1.$$

If either subspace is just the trivial subspace $\{\boldsymbol{0}\}$ (a single point), then the angle is undefined. In other words, this definition applies to subspaces having nonzero dimensions.

If $\boldsymbol{S}$ and $\boldsymbol{T}$ denote orthonormal bases for $\mathcal{S}$ and $\mathcal{T}$, then one can show [4] that $\cos\theta = \|\boldsymbol{S}'\boldsymbol{T}\|_2$ and $\sin\theta = \|\boldsymbol{S}_\perp'\boldsymbol{T}\|_2$ where $\boldsymbol{S}_\perp$ denotes an orthonormal basis for $\mathcal{S}^\perp$ [5].

If $\mathcal{S} \cap \mathcal{T} = \{\boldsymbol{0}\}$, then there is a stronger **Cauchy-Schwarz inequality** [6]:

$$|\langle \boldsymbol{s},\, \boldsymbol{t}\rangle| \leq \gamma \, \|\boldsymbol{s}\|_2 \, \|\boldsymbol{t}\|_2 \,, \ \forall \boldsymbol{s} \in \mathcal{S}, \boldsymbol{t} \in \mathcal{T}, \quad \text{where } 0 \leq \gamma < 1 \text{ depends on } \mathcal{S} \text{ and } \mathcal{T}.$$

One can generalize to examine **angles between flats**. ♦♦

More inner product inequalities (Read)

For the usual **inner product** on $\mathbb{F}^N$:

$$|\langle \boldsymbol{x},\, \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\|_1 \, \|\boldsymbol{y}\|_\infty \,. \tag{5.10}$$

Proof: $|\langle \boldsymbol{x},\, \boldsymbol{y} \rangle| = |\sum_i x_i y_i^*| \leq \sum_i |x_i| \, |y_i| \leq \sum_i |x_i| \, \|\boldsymbol{y}\|_\infty = \|\boldsymbol{x}\|_1 \, \|\boldsymbol{y}\|_\infty \,.$

More generally, if $1/p + 1/q = 1$ and $1 < p, q < \infty$, then **Hölder's inequality** states that

$$|\langle \boldsymbol{x},\, \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\|_p \, \|\boldsymbol{y}\|_q \,, \tag{5.11}$$

again for the usual inner product on $\mathbb{F}^N$.

Using (5.10), the **Frobenius inner product** (5.7) for matrices in $\mathbb{F}^{M \times N}$ satisfies:

$$\mathrm{real}\{\langle \boldsymbol{A},\, \boldsymbol{B} \rangle\} \leq |\langle \boldsymbol{A},\, \boldsymbol{B} \rangle| = |\langle \mathsf{vec}(\boldsymbol{A}),\, \mathsf{vec}(\boldsymbol{B}) \rangle| \leq \|\mathsf{vec}(\boldsymbol{A})\|_1 \, \|\mathsf{vec}(\boldsymbol{B})\|_\infty \,. \tag{5.12}$$

Caution: in general, $\|\mathsf{vec}(\boldsymbol{A})\|_1 \neq |||\boldsymbol{A}|||_1$ and $\|\mathsf{vec}(\boldsymbol{B})\|_\infty \neq |||\boldsymbol{B}|||_\infty \,.$

Challenge. Prove or disprove $|\langle \boldsymbol{A},\, \boldsymbol{B} \rangle| \overset{?}{\leq} |||\boldsymbol{A}|||_1 \, |||\boldsymbol{B}|||_\infty \,.$

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are $N \times N$ square matrices with singular values $\{\sigma_n\}$ and $\{\gamma_n\}$ respectively, then [7]:

$$|\langle \boldsymbol{A},\, \boldsymbol{B} \rangle| = |\mathsf{trace}\{\boldsymbol{A}\boldsymbol{B}'\}| \leq \sum_{n=1}^{N} \sigma_n \gamma_n \leq \sqrt{\left(\sum \sigma_n^2\right) \left(\sum \gamma_n^2\right)} = \|\boldsymbol{A}\|_{\mathrm{F}} \, \|\boldsymbol{B}\|_{\mathrm{F}} \, .$$

Challenge. Generalize the first inequality to include rectangular matrices, or provide a counter-example.

5.2 Matrix norms and operator norms

L§7.4

Also important are **matrix norms** and **operator norms**; roughly speaking these functions quantify "how large" are the elements of a matrix, in different ways.

Define. [1, p. 59] A **matrix norm** on the vector space of matrices $\mathbb{F}^{M\times N}$ is a function $\|\cdot\|$ from $\mathbb{F}^{M\times N}$ to $[0,\infty)$ that satisfies the following properties $\forall \boldsymbol{A}, \boldsymbol{B} \in \mathbb{F}^{M\times N}$:

$\|\boldsymbol{A}\| \geq 0$ (nonnegative)

$\|\boldsymbol{A}\| = 0$ iff $\boldsymbol{A} = \boldsymbol{0}_{M\times N}$ (positive)

$\|\alpha\boldsymbol{A}\| = |\alpha|\,\|\boldsymbol{A}\|$ for all scalars $\alpha \in \mathbb{F}$ in the field (homogeneous)

$\|\boldsymbol{A}+\boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$ (**triangle inequality**)

Because the set of all $M \times N$ matrices $\mathbb{F}^{M\times N}$ is itself a **vector space**, matrix norms are simply vector norms for that space. So at first having a new definition might seem to have modest utility. However, many, *but not all*, matrix norms are **sub-multiplicative**, also called **consistent** [1, p. 61], meaning that they satisfy the following inequality:

$$|||AB||| \leq |||A||| \cdot |||B||| \tag{5.13}$$

This book uses the notation $|||\cdot|||$ to distinguish such **matrix norms** from the ordinary matrix norms $\|\cdot\|$ on the vector space $\mathbb{F}^{M\times N}$ that need not satisfy this extra condition.

Examples of matrix norms

- The **max norm** on $\mathbb{F}^{M \times N}$ is the element-wise maximum: $\|\boldsymbol{A}\|_{\max} \triangleq \max_{i,j} |a_{ij}| = \|\mathrm{vec}(\boldsymbol{A})\|_\infty$.
  This norm is somewhat like the infinity norm for vectors of length MN.
  One can compute it in JULIA using `norm(A, Inf)` after invoking `using LinearAlgebra`.
  Equivalently one may use `norm(A[:], Inf)` because the matrix shape is unimportant for this norm.
  vec(A)
  (However, this differs completely from `opnorm(A, Inf)` that computes $\|\boldsymbol{A}\|_\infty$ described below.)
  The max norm is a matrix norm on the vector space $\mathbb{F}^{M \times N}$ but it does not satisfy the **sub-multiplicative** condition (5.13) so it is of limited use. Most of the norms of interest in SP-DS-ML are sub-multiplicative, so such matrix norms are our primary focus hereafter.

- The **Frobenius norm** (aka **Hilbert-Schmidt norm** or **matrix Eucliden norm**) is defined on $\mathbb{F}^{M \times N}$ by

$$\|\boldsymbol{A}\|_{\mathrm{F}} \triangleq \sqrt{\sum_{m=1}^{M} \sum_{n=1}^{N} |a_{mn}|^2} = \sqrt{\mathrm{trace}\{\boldsymbol{A}'\boldsymbol{A}\}} = \sqrt{\mathrm{trace}\{\boldsymbol{A}\boldsymbol{A}'\}} = \|\mathrm{vec}(\boldsymbol{A})\|_2, \tag{5.14}$$

  and is also called the **Schur norm** and **Schatten 2-norm**. It is a very easy norm to compute.
  The equalities related to trace are an exercise. (HW).
  Practical implementation: `norm(A, 2)` or `norm(A)` or `norm(A[:], 2)` or `norm(A[:])`
  Again, shape of $\boldsymbol{A}$ is unimportant for this norm. vec(A) vec(A)

$A = U_r \Sigma_r V_r'$ $\quad r > 0$

To relate the Frobenius norm of a matrix to its singular values use a **compact SVD**:

$$\|A\|_F = \sqrt{\operatorname{trace}(A'A)} = \sqrt{\operatorname{trace}(V_r \Sigma_r \underbrace{U_r' U_r}_{I} \Sigma_r V_r')} = \sqrt{\operatorname{trace}(\Sigma_r^2 \underbrace{V_r' V_r}_{I})}$$

$$= \sqrt{\operatorname{trace}(\Sigma_r^2)} = \sqrt{\sum_{k=1}^{r} \sigma_k^2} = \|\Sigma_r\|_F = \sqrt{\operatorname{trace}(\Sigma_r^2)} = \sqrt{\operatorname{tr}(\Sigma\Sigma')}$$

$$= \sqrt{\sum_{k=1}^{\min(M,N)} \sigma_k^2}$$

This norm is invariant to unitary transformations [8, p. 442], because of the trace property (1.32) (see p. 5.10).

From a vector perspective, this norm is induced by the **Frobenius inner product** on $\mathbb{F}^{M \times N}$. As a matrix ♦♦ norm however, it is not induced by any vector norm on $\mathbb{F}^N$ (see next page) [9], but nevertheless it is **compatible** with the Euclidean vector norm because

$$\boxed{\|Ax\|_2 \leq \|A\|_F \|x\|_2 .} \tag{5.15}$$

However, this upper bound is not tight in general. (It is tight for rank-1 matrices only.)

By combining (5.15) with the definition of matrix multiplication, one can show easily that the Frobenius norm is **sub-multiplicative** [10, p. 291].

$$u v' = \frac{u}{\|u\|} \underbrace{\|u\| \cdot \|v\|}_{\sigma_1} \frac{v'}{\|v\|} \qquad r = 1$$

5. What is the Frobenius norm of the **outer product** $\|uv'\|_F$ for $u \in \mathbb{F}^M$, $v \in \mathbb{F}^N$?

A: $\|u\|_2 \|v\|_2$ B: $\sqrt{\|u\|_2 \|v\|_2}$ C: $|u'v|^2$ D: $|u'v|$ E: None of these. ??

$$\|\underbrace{u v'}_{A}\|_F = \sqrt{\operatorname{trace}(\underbrace{v u'}_{A'} \underbrace{u v'}_{A})} = \sqrt{\operatorname{trace}(u'u \, v'v)} = \|u\|_2 \cdot \|v\|_2$$

$\ell_{p,q}$ norms (Read)

For certain signal processing problems involving **group sparsity** [11, 12], the following family of $\ell_{p,q}$ **matrix norms** is useful: ♦♦

$$\|A\|_{p,q} \triangleq \left(\sum_{n=1}^{N} \left(\|A_{:,n}\|_p \right)^q \right)^{1/q} = \left(\sum_{n=1}^{N} \left(\sum_{m=1}^{M} |a_{m,n}|^p \right)^{q/p} \right)^{1/q}.$$

This family considers a $M \times N$ matrix A as a collection of N columns of length M.

A particularly popular special case for group sparsity problems is

$$\|A\|_{1,2} = \left(\sum_{n=1}^{N} \left(\|A_{:,n}\|_1 \right)^2 \right)^{1/2}.$$

Note that in general $\|A\|_{p,p} = \|\mathrm{vec}(A)\|_p$ and specifically $\|A\|_{2,2} = \|A\|_{\mathrm{F}}$.

Challenge. Determine which of the $\ell_{p,q}$ norms are **sub-multiplicative**.

Induced matrix norms

Define. If $\|\cdot\|$ is any vector norm that is suitable for both $\mathbb{F}^N$ and $\mathbb{F}^M$, then a matrix norm for $\mathbb{F}^{M\times N}$ is:

$$\|\boldsymbol{A}\| \triangleq \max_{x \in \mathbb{F}^N : \|x\|=1} \|Ax\| = \max_{x \neq 0_N} \frac{\|Ax\|}{\|x\|} \tag{5.16}$$

which is called an **operator norm** (because now $\boldsymbol{A}$ acts as an operation).
We say such a matrix norm $\|\cdot\|$ is **induced** by the vector norm $\|\cdot\|$.

By construction, if $\|\cdot\|$ is a matrix norm induced by vector norm $\|\cdot\|$, then:

tight inequality

$$\boxed{\|Ax\| \leq |||A||| \, \|x\|} \quad \forall x \in \mathbb{F}^N \tag{5.17}$$

Importantly, the **sub-multiplicative** property (5.13) holds for any **induced norm** provided the number of columns of $\boldsymbol{A}$ matches the number of rows of $\boldsymbol{B}$. This fact follows readily from the definition (5.16) and the property (5.17) because

$$\|\boldsymbol{A}\boldsymbol{B}\| = \max_{\boldsymbol{x} : \|\boldsymbol{x}\|=1} \|\boldsymbol{A}\boldsymbol{B}\boldsymbol{x}\| \leq \max_{\boldsymbol{x} : \|\boldsymbol{x}\|=1} \|\boldsymbol{A}\| \, \|\boldsymbol{B}\boldsymbol{x}\| = \|\boldsymbol{A}\| \, \|\boldsymbol{B}\| \,.$$

Example. The most important matrix norms (**operator norms**) are induced by the vector norm $\|\cdot\|_p$, *i.e.*,

$$\|A\|_p = \max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}. \tag{5.18}$$

- The **spectral norm** $|||\cdot|||_2$, often denoted simply $|||\cdot|||$, is defined on $\mathbb{F}^{M \times N}$ by (5.18) with $p = 2$. This is the matrix norm induced by the Euclidean vector norm. As shown on p. 2.36:

$$|||A|||_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \max\left\{\sqrt{\lambda} \,:\, \lambda \in \mathrm{eig}\{A'A\}\right\} = \sigma_1(A)$$

ch2

- The **maximum row sum matrix norm** is defined on $\mathbb{F}^{M \times N}$ by

$$|||A|||_\infty \triangleq \max_{1 \le i \le M} \sum_{j=1}^{N} |a_{ij}| \,. \tag{5.19}$$

It is induced by the ℓ_∞ vector norm. It differs from the max norm defined above! Here the shape matters!

Proof:

$$\begin{aligned}\|A\|_\infty &= \max_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_{x \neq 0} \frac{\max_{m=1,\ldots,M} |[Ax]_m|}{\|x\|_\infty} = \max_{x \neq 0} \frac{\max_{m=1,\ldots,M} \left|\sum_{n=1}^N a_{mn} x_n\right|}{\|x\|_\infty} \\ &\leq \max_{x \neq 0} \frac{\max_{m=1,\ldots,M} \sum_{n=1}^N |a_{mn}| \, |x_n|}{\|x\|_\infty} \leq \max_{x \neq 0} \frac{\max_{m=1,\ldots,M} \sum_{n=1}^N |a_{mn}| \, \|x\|_\infty}{\|x\|_\infty} = \max_{1 \leq m \leq M} \sum_{n=1}^N |a_{mn}| .\end{aligned}$$

- The **maximum column sum matrix norm** is defined on $\mathbb{F}^{M \times N}$ by

$$\|A\|_1 \triangleq \max_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1} = \max_{1 \leq j \leq N} \sum_{i=1}^M |a_{ij}| . \tag{5.20}$$

It is induced by the ℓ_1 vector norm. Note that $\|A\|_1 = \|A'\|_\infty$.

Norms defined in terms of singular values

Here are three important norms used in modern signal processing problems.

- The **nuclear norm**, sometimes called the **trace norm** [1, p. 60], is the sum of the singular values:

$$\boxed{\|\boldsymbol{A}\|_* \triangleq \sum_{k=1}^{\min(M,N)} \sigma_k} = \mathsf{trace}\{(\boldsymbol{A}'\boldsymbol{A})^{1/2}\} = \mathsf{trace}\{(\boldsymbol{A}\boldsymbol{A}')^{1/2}\},$$

where **matrix square root** is discussed in Ch. 8.

- For $1 \leq p \leq \infty$, the **Schatten p-norm** of a $M \times N$ matrix is defined using the ℓ_p norm of its singular values:

$$\|\boldsymbol{A}\|_{S,p} = \left(\sum_{k=1}^{\min(M,N)} \sigma_k^p\right)^{1/p} = \left(\sum_{k=1}^{\mathsf{rank}(\boldsymbol{A})} \sigma_k^p\right)^{1/p}.$$

- The **Ky-Fan K-norm** is the sum of the first $1 \leq K \leq \min(M, N)$ singular values of a matrix:

$$\|\boldsymbol{A}\|_{\text{Ky-Fan,K}} = \sum_{k=1}^{K} \sigma_k(\boldsymbol{A}).$$

For a PCA generalization that uses this norm see [13].
For a Schatten 4-norm used in a data science application see [14]. For an imaging application that uses $p = 0.9$ (not a norm, but useful still in practice) see [15].

- For (complicated!) proofs that these are in fact norms (*i.e.*, satisfy the triangle inequality), see [16, p.91].
- All of these three norms (nuclear, Schatten, and Ky-Fan) are **sub-multiplicative** [16, p.94].
  In particular, this means that $\|\boldsymbol{B}\boldsymbol{A}\|_* \leq \|\boldsymbol{B}\|_* \|\boldsymbol{A}\|_*$.
  However, apparently there is a tighter inequality: $\|\boldsymbol{B}\boldsymbol{A}\|_* \leq \|\boldsymbol{B}\|_2 \|\boldsymbol{A}\|_*$.

Challenge. Prove (or disprove) that inequality.

Von Neuman's trace inequality states that if $\boldsymbol{A}$ and $\boldsymbol{B}$ are both in $\mathbb{F}^{N\times N}$, then ♦♦

$$|\text{trace}\{\boldsymbol{A}\boldsymbol{B}\}| \leq \sum_{n=1}^{N} \sigma_n(\boldsymbol{A})\sigma_n(\boldsymbol{B}) \leq \sum_{n=1}^{N} \sigma_n(\boldsymbol{A})\sigma_1(\boldsymbol{B}) = \|\boldsymbol{B}\|_2 \|\boldsymbol{A}\|_*.$$

Thus

$$\|\boldsymbol{B}\|_2 \leq 1 \Longrightarrow |\text{trace}\{\boldsymbol{A}\boldsymbol{B}\}| \leq \|\boldsymbol{A}\|_* \Longrightarrow \sup_{\boldsymbol{B}\,:\,\|\boldsymbol{B}\|_2 \leq 1} |\text{trace}\{\boldsymbol{A}\boldsymbol{B}\}| \leq \|\boldsymbol{A}\|_*.$$

On the other hand, if $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}'$ and we define $\boldsymbol{B} = \boldsymbol{V}\boldsymbol{U}'$, then $|\text{trace}\{\boldsymbol{A}\boldsymbol{B}\}| = |\text{trace}\{\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}'\boldsymbol{V}\boldsymbol{U}'\}| = \|\boldsymbol{A}\|_*$, so $\sup_{\boldsymbol{B}\,:\,\|\boldsymbol{B}\|_2 \leq 1} |\text{trace}\{\boldsymbol{A}\boldsymbol{B}\}| \geq \|\boldsymbol{A}\|_*$. Combining yields this **nuclear norm** equality:

$$\|\boldsymbol{A}\|_* = \sup_{\boldsymbol{B}\,:\,\|\boldsymbol{B}\|_2 \leq 1} |\text{trace}\{\boldsymbol{A}\boldsymbol{B}\}|.$$

Relationships between these norms:

- Nuclear norm:

$$\|A\|_* = \|A\|_{S,1} = \|A\|_{\mathrm{Ky-Fan,min(M,N)}}$$

- Spectral norm:

$$\|A\|_2 = \sigma_1(A) = \|A\|_{S,\infty} = \|A\|_{\mathrm{Ky-Fan,1}}$$

- Frobenius norm:

$$\|A\|_{\mathrm{F}} = \|A\|_{S,2}$$

Exercise. Relate $\|A\|_{\mathrm{F}}$ to a Ky-Fan norm and to a nuclear norm involving A. $\|A\|_{\mathrm{F}}^2 = \|A'A\|_{\mathrm{Ky-Fan,min(M,N)}} = \|A'A\|_*$

Challenge. Prove whether the **Schatten p-norm** $\|\cdot\|_{S,p}$ is or is not an **induced norm** for $1 \le p < \infty$. ♦♦

Exercise. Define a matrix norm that unifies all of the matrix norms defined here in terms of singular values.

$$|||A|||_{K,p} = \left(\sum_{k=1}^{K} \sigma_k^p\right)^{1/p}$$

$K = \min(M,N) \Rightarrow$ Schatten p-norm

Practical implementation

JULIA commands (after invoking `using LinearAlgebra`) for some of these norms are as follows:

- $\|A\|_1$ `opnorm(A, 1)`
- $\|A\|_2$ `opnorm(A, 2)` or just `opnorm(A)` = maximum(svdvals(A)) = svdvals(A)[1]
- $\|A\|_\infty$ `opnorm(A, Inf)` ($\|A\|_{\max}$ is `norm(A, Inf)`)
- $\|A\|_*$ `sum(svdvals(A))` or `sum(svd(A).S)`

Preview of practical use of such norms:

- $\|A\|_2$ is relatively expensive to compute because it requires an SVD of A, whereas $\|A\|_1$ and $\|A\|_\infty$ are easy to compute. See p. 5.40
- $\|A\|_*$ is useful as a convex relaxation of $\operatorname{rank}(A)$; see Ch. 6.

$$A = \begin{bmatrix} 1 & 1 & 1 \\ -3 & -3 & -3 \end{bmatrix}$$

Examples

6. For $A = [1\ -3]'\,[1\ 1\ 1]$, what is `norm(A, Inf)` ?

A: 2 B: 3 C: 6 D: 9 E: None of these ??

7. For $A = [1\ -3]'\,[1\ 1\ 1]$, what is `opnorm(A, Inf)` ?

A: 2 B: 3 C: 6 D: 9 E: None of these ??

8. For $A = \underbrace{[1\ -3]'}_{U}\,\underbrace{[1\ 1\ 1]}_{V'}$, what is `opnorm(A, 2)` ?

A: 2	B: 3	C: 6	D: 9	E: None of these	**??**

$\sigma_1 = \|u\| \cdot \|v\| = \sqrt{30}$

Properties of matrix norms

All matrix norms are also **equivalent** (to within constants that depend on the matrix dimensions).
See [1, p. 61] for inequalities relating various matrix norms.

Example. (HW) $$A \in \mathbb{F}^{M \times N} \implies \|A\|_1 \leq \sqrt{M}\,\|A\|_2 .$$

Example. To relate the **spectral norm** and **nuclear norm** for a matrix A having rank r:

$$\|A\|_* = \sum_{k=1}^{r} \sigma_k \leq \sum_{k=1}^{r} \sigma_1 = r\,\sigma_1 = r\,|||A|||_*$$

$$\|A\|_2 = \sigma_1 \leq \sum_{k=1}^{r} \sigma_k = |||A|||_*$$

Combining yields the following inequalities:

$$\frac{1}{r}\,|||A|||_* \leq |||A|||_2 \leq |||A|||_* \leq r\,|||A|||_2$$

To express it in way that depends on the norm only (not r, which is a property of a specific matrix):

$$r \leq \min(M, N)$$

2021-10-12

$$\frac{1}{\min(M,N)} |||A|||_* \leq |||A|||_2 \leq |||A|||_* \leq \min(M,N) \, |||A|||_2$$

Singular value inequalities (Read)

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are both $M \times N$ matrices, then the **triangle inequality** for the spectral norm is

$$\|\boldsymbol{A} + \boldsymbol{B}\|_2 \leq \|\boldsymbol{A}\|_2 + \|\boldsymbol{B}\|_2 \, ,$$

or equivalently, in terms of singular values:

$$\sigma_1(\boldsymbol{A} + \boldsymbol{B}) \leq \sigma_1(\boldsymbol{A}) + \sigma_1(\boldsymbol{B}).$$

The following inequality provides a generalization to other singular values [17]: ♦♦

$$\sigma_{m+n+1}(\boldsymbol{A} + \boldsymbol{B}) \leq \sigma_{m+1}(\boldsymbol{A}) + \sigma_{n+1}(\boldsymbol{B}), \qquad \begin{array}{c} 0 \leq m, n \leq \min(M, N) - 1 \\ m + n + 1 \leq \min(M, N). \end{array}$$

Example. Considering $m + n + 1 = 3$ we have

$$\begin{aligned}
\sigma_3(\boldsymbol{A} + \boldsymbol{B}) &\leq \sigma_3(\boldsymbol{A}) + \sigma_1(\boldsymbol{B}) \quad (m = 2, n = 0) \\
\sigma_3(\boldsymbol{A} + \boldsymbol{B}) &\leq \sigma_2(\boldsymbol{A}) + \sigma_2(\boldsymbol{B}) \quad (m = 1, n = 1) \\
\sigma_3(\boldsymbol{A} + \boldsymbol{B}) &\leq \sigma_1(\boldsymbol{A}) + \sigma_3(\boldsymbol{B}) \quad (m = 0, n = 2) \\
\Longrightarrow \sigma_3(\boldsymbol{A} + \boldsymbol{B}) &\leq \min(\sigma_3(\boldsymbol{A}) + \sigma_1(\boldsymbol{B}), \sigma_2(\boldsymbol{A}) + \sigma_2(\boldsymbol{B}), \sigma_1(\boldsymbol{A}) + \sigma_3(\boldsymbol{B})).
\end{aligned}$$

Such inequalities are used far less often than the triangle inequality.

Unitarily invariant matrix norms

Define. A matrix norm $\|\cdot\|$ on $\mathbb{F}^{M\times N}$ is called **unitarily invariant** iff for all unitary matrices $\boldsymbol{U} \in \mathbb{F}^{M\times M}$ and $\boldsymbol{V} \in \mathbb{F}^{N\times N}$:

$$\|\boldsymbol{U}\boldsymbol{A}\boldsymbol{V}\| = \|\boldsymbol{A}\|, \quad \forall \boldsymbol{A} \in \mathbb{F}^{M\times N}.$$

Theorem.

- The **spectral norm** $\|\boldsymbol{A}\|_2$ is unitarily invariant
- Any **Schatten p-norm** $\|\boldsymbol{A}\|_{S,p}$ is unitarily invariant
  Proof sketch: unitary matrix rotations do not change singular values.
- The **Frobenius norm** $\|\boldsymbol{A}\|_{\mathrm{F}}$ is unitarily invariant
  Proof for Frobenius case:

$$\|\underbrace{\boldsymbol{U}\boldsymbol{A}\boldsymbol{V}}_{M}\|_{\mathrm{F}}^2 = \mathrm{trace}(\overbrace{UAV}^{M}\,\overbrace{V'A'U'}^{M'}) = \mathrm{trace}(AA'\underbrace{U'U}_{I})$$

$$= \mathrm{trace}(AA') = \|A\|_F^2 = \sum_{k=1}^{r} \sigma_k^2$$

(We could also prove it using singular values.)

The Frobenius norm has an even more general invariance. If $\boldsymbol{U}$ has M orthonormal columns and $\boldsymbol{Q}$ has N orthonormal rows, then by the same proof $\|\boldsymbol{U}\boldsymbol{A}\boldsymbol{Q}\|_{\mathrm{F}} = \|\boldsymbol{A}\|_{\mathrm{F}}$ for any $\boldsymbol{A} \in \mathbb{F}^{M\times N}$.

Fact. Every **unitarily invariant** matrix norm is **sub-multiplicative**. (See [16, p.94] for complicated proof.)

$$\|AB\| \leq \|A\|\|B\|$$

Fact. If $\boldsymbol{A}$ and $\boldsymbol{B}$ are **positive semi-definite** matrices, then for every unitarily invariant norm [18]:

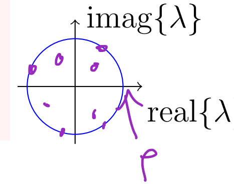
$$2 \|\boldsymbol{AB}\| \leq \left\|\boldsymbol{A}^2 + \boldsymbol{B}^2\right\| \text{ and } 4 \|\boldsymbol{AB}\| \leq \left\|(\boldsymbol{A}+\boldsymbol{B})^2\right\|.$$

These properties generalize the **arithmetic–geometric mean inequality** to PSD matrices.

Spectral radius

Define. For any square matrix, the **spectral radius** is the maximum absolute eigenvalue:

$$\boldsymbol{A} \in \mathbb{F}^{N \times N} \Longrightarrow \rho(\boldsymbol{A}) \triangleq \max_i |\lambda_i(\boldsymbol{A})|$$



- By construction, $|\lambda_i(\boldsymbol{A})| \leq \rho(\boldsymbol{A})$ so all eigenvalues lie within a disk in the complex plane of radius $\rho(\boldsymbol{A})$, hence the name.
- In general, $\rho(\boldsymbol{A})$ is *not* a **matrix norm** and $\|\boldsymbol{A}\boldsymbol{x}\| \not\leq \rho(\boldsymbol{A}) \|\boldsymbol{x}\|$. Consider $\boldsymbol{A} = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$. $\rho(A)=0$ $\|A\|_2 \neq 0$
- However, if $\boldsymbol{A}$ is **normal**, then recall from (2.6) that we if order its eigenvalues in decreasing order of their absolute values, then we can relate its unitary eigendecomposition to an SVD as follows:

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}' = \sum_{n=1}^{N} \lambda_n \boldsymbol{v}_n \boldsymbol{v}_n' = \sum_{n=1}^{N} \underbrace{|\lambda_n|}_{\sigma_n} \underbrace{\operatorname{sign}(\lambda_n)\, \boldsymbol{v}_n}_{\boldsymbol{u}_n} \boldsymbol{v}_n'.$$

  - Thus if $\boldsymbol{A}$ is **normal** (*e.g.*, $\boldsymbol{A} = \boldsymbol{A}'$) then $\boxed{\rho(\boldsymbol{A}) = \sigma_1(\boldsymbol{A}) = \|\boldsymbol{A}\|_2}$, so $\|\boldsymbol{A}\boldsymbol{x}\|_2 \leq \rho(\boldsymbol{A}) \|\boldsymbol{x}\|_2$.
  - Furthermore, $\boldsymbol{A}$ **normal** $\Longrightarrow |\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}| \leq \rho(\boldsymbol{A}) \|\boldsymbol{x}\|_2^2$ because, applying the Cauchy-Shwarz inequality:

$$|\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}| = |\langle \boldsymbol{A}\boldsymbol{x},\, \boldsymbol{x}\rangle| \leq \|\boldsymbol{A}\boldsymbol{x}\|_2 \|\boldsymbol{x}\|_2 \leq \|\boldsymbol{A}\|_2 \|\boldsymbol{x}\|_2^2 = \rho(\boldsymbol{A}) \|\boldsymbol{x}\|_2^2 .$$

If $\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and $\boldsymbol{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ then $\rho(\boldsymbol{A}) = 1$ but $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = 6 > \|\boldsymbol{x}\|_2^2 = 5$.
So $\boldsymbol{A}$ being diagonalizable is insufficient to conclude that $|\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}| \leq \rho(\boldsymbol{A}) \|\boldsymbol{x}\|_2^2$.

- If $\|\cdot\|$ is any **induced matrix norm** on $\mathbb{F}^{N\times N}$ and if $\boldsymbol{A} \in \mathbb{F}^{N\times N}$, then

$$\rho(\boldsymbol{A}) \leq \|\boldsymbol{A}\| . \tag{5.21}$$

Proof. If $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$, then $|\lambda| \|\boldsymbol{v}\| = \|\lambda\boldsymbol{v}\| = \|\boldsymbol{A}\boldsymbol{v}\| \leq \|\boldsymbol{A}\| \|\boldsymbol{v}\|$. Dividing by $\|\boldsymbol{v}\|$, which is fine because $\boldsymbol{v} \neq \boldsymbol{0}$, yields $|\lambda| \leq \|\boldsymbol{A}\|$. This inequality holds for all eigenvalues, including the one with maximum magnitude. □

The same proof shows that $\rho(\boldsymbol{A}) \leq \|\boldsymbol{A}\|_{\mathrm{F}}$ because of the compatible property (5.15).

- If $\boldsymbol{A} \in \mathbb{F}^{N\times N}$, then $\lim_{k\to\infty} \boldsymbol{A}^k = \boldsymbol{0}$ if and only if $\rho(\boldsymbol{A}) < 1$.
This property is particularly important for analyzing the convergence of iterative algorithms, including training **recurrent neural networks** [19]. (*cf.* HW)
- For any $\boldsymbol{A} \in \mathbb{F}^{N\times N}$, the **spectral radius** is an infimum of all **induced matrix norms**: ♦♦

$$\rho(\boldsymbol{A}) = \inf \{ \|\boldsymbol{A}\| \ : \ \|\cdot\| \text{ is an induced matrix norm} \} .$$

- **Gelfand's formula** for any **induced matrix norm** $\|\cdot\|$ for a square matrix $\boldsymbol{A}$ is: ♦♦

$$\rho(\boldsymbol{A}) = \lim_{k\to\infty} \left\| \boldsymbol{A}^k \right\|^{1/k} . \tag{5.22}$$

 $\rho(A) = \max_i |\lambda_i(A)|$

9. Which equality (if any) correctly relates a singular value and a spectral radius for any general matrix $\boldsymbol{A} \in \mathbb{F}^{M \times N}$?

A: $\sigma_1(\boldsymbol{A}) \stackrel{?}{=} |\rho(\boldsymbol{A})|$

B: $\sigma_1(\boldsymbol{A}) \stackrel{?}{=} \rho^2(\boldsymbol{A})$

C: $\sigma_1(\boldsymbol{A}) \stackrel{?}{=} \rho(\boldsymbol{A}'\boldsymbol{A})$

D: $\sigma_1(\boldsymbol{A}) \stackrel{?}{=} \sqrt{\rho(\boldsymbol{A}'\boldsymbol{A})}$ $\qquad \rho(A'A) = \max_i |\lambda_i(A'A)| = \max_i \lambda_i(A'A)$

E: None of these.

$\boxed{??}$ $\qquad = \max_i \sigma_i(A'A) = \sigma_1(A'A) = \sigma_1^2(A)$

(Read)

We have seen that $\boldsymbol{A}$ **normal** $\Longrightarrow \rho(\boldsymbol{A}) = \sigma_1(\boldsymbol{A})$, but the converse is not true.

Example. Consider $\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix}$ for which $\rho(\boldsymbol{A}) = \sigma_1(\boldsymbol{A}) = \sqrt{2}$, but $\boldsymbol{A}$ is not normal.

$AA' \neq A'A$

Challenge. Find a 2×2 matrix example where $\rho(\boldsymbol{A}) = \sigma_1(\boldsymbol{A})$ $\boldsymbol{A}$ is not normal, or prove that none exist.

Practical step size for gradient descent

The **gradient descent** (**GD**) method $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \mu \boldsymbol{A}'(\boldsymbol{A}\boldsymbol{x}_k - \boldsymbol{y})$ for solving a linear least-squares problem $\arg\min_{\boldsymbol{x}} \frac{1}{2} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|$ **converges** (for any $\boldsymbol{x}_0$; see Ch. 8) iff $\rho(\boldsymbol{I} - \mu \boldsymbol{A}'\boldsymbol{A}) < 1$, *i.e.*, iff the step size μ satisfies

$$\boxed{0 < \mu < \frac{2}{\sigma_1^2(\boldsymbol{A})}.} \tag{5.23}$$

However, computing the spectral norm $\sigma_1(\boldsymbol{A}) = \|\boldsymbol{A}\|_2$ is expensive for large problems.

A more practical alternative is to use the inequality (5.21) to find an upper bound on the spectral norm:

$$\sigma_1^2(\boldsymbol{A}) = \sigma_1(A'A) = \|A'A\|_2 = \rho(A'A) \leq \|A'A\|_1$$

Similarly, $\sigma_1^2(\boldsymbol{A}) \leq \|\boldsymbol{A}\boldsymbol{A}'\|_1$, so combining yields the upper bound

$$\sigma_1^2(\boldsymbol{A}) \leq \min(\|\boldsymbol{A}'\boldsymbol{A}\|_1, \|\boldsymbol{A}\boldsymbol{A}'\|_1).$$

Thus, choosing

$$0 < \mu < \frac{2}{\min(\|\boldsymbol{A}'\boldsymbol{A}\|_1, \|\boldsymbol{A}\boldsymbol{A}'\|_1)}$$

would provide a step size that lies in the proper interval in (5.23).

Now if $\boldsymbol{A}$ is a $M \times N$ matrix, then computing $\boldsymbol{A}'\boldsymbol{A}$ would require N^2M multiplies whereas $\boldsymbol{A}\boldsymbol{A}'$ would require M^2N multiplies. These are both expensive if M and N are large. In fact, they have the same order

$O(MN^2)$ (for $M \geq N$) as computing an SVD of $\boldsymbol{A}$. To avoid matrix multiplication, we can use the following upper bound:

$$\sigma_1^2(\boldsymbol{A}) = \sigma_1(\boldsymbol{A}'\boldsymbol{A}) = \|\boldsymbol{A}'\boldsymbol{A}\|_2 = \rho(\boldsymbol{A}'\boldsymbol{A}) \leq \|\boldsymbol{A}'\boldsymbol{A}\|_1 \leq |||A'|||_1 \cdot |||A|||_1 = |||A|||_\infty \cdot |||A|||_1$$

because the matrix 1-norm is sub-multiplicative. Thus, choosing

$$\boxed{0 < \mu < \frac{2}{\|\boldsymbol{A}\|_\infty \|\boldsymbol{A}\|_1}} \tag{5.24}$$

is a valid step size that lies in the proper interval in (5.23), and hence ensures GD converges (*cf.* `lsgd` and `lsngd`).

It is much easier to compute $\|\boldsymbol{A}\|_\infty$ and $\|\boldsymbol{A}\|_1$ than $\|\boldsymbol{A}\|_2$.

$M(N-1)$

10. Approximately how many additions are needed to compute $\|\boldsymbol{A}\|_1$?

A: M B: N C: $\min(M, N)$ D: $\max(M, N)$ E: MN **??**

The drawback of (5.24) is that it may be a loose upper bound. If $\boldsymbol{A}$ is the $N \times N$ unitary DFT matrix discussed in Ch. 7, then $\|\boldsymbol{A}\|_2 = 1$ but $\|\boldsymbol{A}\|_\infty = \|\boldsymbol{A}\|_1 = \sqrt{N}$.

An exercise (HW) shows that $\|\boldsymbol{A}\|_\infty \leq \sqrt{N} \|\boldsymbol{A}\|_2$ for a $M \times N$ matrix $\boldsymbol{A}$, so this unitary DFT example achieves that worst-case upper bound.

5.3 Convergence of sequences of vectors and matrices

(Read)

Later chapters discuss iterative optimization algorithms and analyze when sequences produced by such algorithms **converge**. This is another topic involving **vector norms** and **matrix norms**.

Recall the following definition for convergence of a sequence of numbers.

Define. We say a sequence of (possibly complex) numbers $\{x_k\}$ **converges** to a limit x_* iff $|x_k - x_*| \to 0$ as $k \to \infty$, where $|\cdot|$ denotes absolute value (or complex magnitude more generally). Specifically,

$$\forall \epsilon > 0, \quad \exists N_\epsilon \in \mathbb{N} \text{ s.t. } |x_k - x_*| < \epsilon \quad \forall k \geq N_\epsilon.$$

We now define convergence of a sequence of vectors or matrices by using a **norm** to quantify distance, then relating convergence of vectors to that of a sequence of scalars.

Define. We say a sequence of vectors $\{\boldsymbol{x}_k\}$ in a vector space $\mathcal{V}$ **converges** to a limit $\boldsymbol{x}_* \in \mathcal{V}$ iff $\|\boldsymbol{x}_k - \boldsymbol{x}_*\| \to 0$ for some norm $\|\cdot\|$ as $k \to \infty$. Specifically,

$$\forall \epsilon > 0, \quad \exists N_\epsilon \in \mathbb{N} \text{ s.t. } \|\boldsymbol{x}_k - \boldsymbol{x}_*\| < \epsilon \quad \forall k \geq N_\epsilon$$

Often we write $\boldsymbol{x}_k \to \boldsymbol{x}_*$ as a shorthand for $\|\boldsymbol{x}_k - \boldsymbol{x}_*\| \to 0$.

A matrix is simply a point in a vector space of matrices so we use essentially the same definition of convergence of a sequence of matrices.

Define. We say a sequence of matrices $\{\boldsymbol{X}_k\}$ (in a vector space $\mathcal{V}$ of matrices) **converges** to a limit $\boldsymbol{X}_* \in \mathcal{V}$ iff $\|\boldsymbol{X}_k - \boldsymbol{X}_*\| \to 0$ for some (matrix) norm $\|\cdot\|$ as $k \to \infty$. Specifically,

$$\forall \epsilon > 0, \quad \exists N_\epsilon \in \mathbb{N} \text{ s.t. } \|\boldsymbol{X}_k - \boldsymbol{X}_*\| < \epsilon \quad \forall k \geq N_\epsilon$$

Example. Consider (for simplicity) the sequence of **diagonal** matrices $\{\boldsymbol{D}_k\}$ defined by

$$\boldsymbol{D}_k = \begin{bmatrix} 3 + 2^{-k} & 0 \\ 0 & (-1)^k/k^2 \end{bmatrix}.$$

This sequence converges to the limit $\boldsymbol{D}_* = \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}$ because

$$\|\boldsymbol{D}_k - \boldsymbol{D}_*\|_{\mathrm{F}} = \left\| \begin{bmatrix} 2^{-k} & 0 \\ 0 & (-1)^k/k^2 \end{bmatrix} \right\|_{\mathrm{F}} = \sqrt{4^{-k} + 1/k^4} \to 0.$$

Example. For a square matrix $\boldsymbol{A}$, define the partial sum of powers $\boldsymbol{S}_k \triangleq \sum_{j=0}^{k} \boldsymbol{A}^j$. If $\|\boldsymbol{A}\|_2 < 1$, then one can show that $\boldsymbol{I} - \boldsymbol{A}$ is invertible and the matrix sequence $\{\boldsymbol{S}_k\}$ converges to the **Neumann series**: $\sum_{j=0}^{\infty} \boldsymbol{A}^j = (\boldsymbol{I} - \boldsymbol{A})^{-1}$. See [20] for a recent use in neural networks.

5.4 Generalized inverse of a matrix

(Read)

The **Moore-Penrose pseudoinverse** defined on p. 4.25 is just one (particularly important) type of **generalized inverse** of a matrix. This section uses the **Frobenius norm** to characterize the **Moore-Penrose pseudoinverse**.

Define. A matrix $\boldsymbol{G} \in \mathbb{F}^{N \times M}$ is a **generalized inverse** of a matrix $\boldsymbol{A} \in \mathbb{F}^{M \times N}$ iff $\boldsymbol{A}\boldsymbol{G}\boldsymbol{A} = \boldsymbol{A}$.

- If $\boldsymbol{A}$ has full column rank, then $\boldsymbol{A}'\boldsymbol{A}$ is invertible, so multiplying both sizes of $\boldsymbol{A}\boldsymbol{G}\boldsymbol{A} = \boldsymbol{A}$ on the left by $\boldsymbol{A}^{+} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'$ yields that $\boldsymbol{G}$ is a generalized inverse of such an $\boldsymbol{A}$ iff $\boldsymbol{G}\boldsymbol{A} = \boldsymbol{I}_N$, *i.e.*, iff $\boldsymbol{G}$ is a **left inverse** of $\boldsymbol{A}$.
- Conversely, if $\boldsymbol{A}$ has full row rank, then $\boldsymbol{A}\boldsymbol{A}'$ is invertible and $\boldsymbol{G}$ is a generalized inverse of such an $\boldsymbol{A}$ iff $\boldsymbol{A}\boldsymbol{G} = \boldsymbol{I}_M$, *i.e.*, iff $\boldsymbol{G}$ is a **right inverse** of $\boldsymbol{A}$.

Considering an **SVD** $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}'$, one can verify from the definition that every generalized inverse of $\boldsymbol{A}$ has the form

$$\boldsymbol{G} = \boldsymbol{V} \begin{bmatrix} \boldsymbol{\Sigma}_r^{-1} & \boldsymbol{S}_2 \\ \boldsymbol{S}_3 & \boldsymbol{S}_4 \end{bmatrix} \boldsymbol{U}',$$

where the matrices $\boldsymbol{S}_2$, $\boldsymbol{S}_3$, $\boldsymbol{S}_4$ have certain sizes but otherwise have completely arbitrary values. In other words, the (very general!) set of generalized inverses $\mathcal{G}_{\boldsymbol{A}}$ of a $M \times N$ matrix $\boldsymbol{A}$ is a **linear variety** in the vector space of $N \times M$ matrices.

Exercise. Determine the sizes of $\boldsymbol{S}_2$, $\boldsymbol{S}_3$, $\boldsymbol{S}_4$. $\boldsymbol{S}_2 \in \mathbb{F}^{r \times N-r}$, $\boldsymbol{S}_3 \in \mathbb{F}^{M-r \times r}$, $\boldsymbol{S}_4 \in \mathbb{F}^{M-r \times N-r}$

One can devise many ways to choose a specific generalized inverse from the set $\mathcal{G}_{\boldsymbol{A}}$.

Minimum Frobenius norm generalized inverse

A simple way is to choose the generalized inverse having the smallest **Frobenius norm**. This solution turns out to be simply the pseudo-inverse of $\boldsymbol{A}$:

$$\operatorname*{arg\,min}_{\boldsymbol{G} \in \mathcal{G}_{\boldsymbol{A}}} \|\boldsymbol{G}\|_{\mathrm{F}} = \boldsymbol{A}^+.$$

Proof: $\boldsymbol{G} \in \mathcal{G}_{\boldsymbol{A}} \implies \boldsymbol{G} = \boldsymbol{V} \begin{bmatrix} \boldsymbol{\Sigma}_r^{-1} & \boldsymbol{S}_2 \\ \boldsymbol{S}_3 & \boldsymbol{S}_4 \end{bmatrix} \boldsymbol{U}' \implies \|\boldsymbol{G}\|_{\mathrm{F}} = \left\| \begin{bmatrix} \boldsymbol{\Sigma}_r^{-1} & \boldsymbol{S}_2 \\ \boldsymbol{S}_3 & \boldsymbol{S}_4 \end{bmatrix} \right\|_{\mathrm{F}}$ because the Frobenius norm is **unitarily invariant**. Because $\left\| \begin{bmatrix} \boldsymbol{\Sigma}_r^{-1} & \boldsymbol{S}_2 \\ \boldsymbol{S}_3 & \boldsymbol{S}_4 \end{bmatrix} \right\|_{\mathrm{F}}^2 = \|\boldsymbol{\Sigma}_r^{-1}\|_{\mathrm{F}}^2 + \|\boldsymbol{S}_2\|_{\mathrm{F}}^2 + \|\boldsymbol{S}_3\|_{\mathrm{F}}^2 + \|\boldsymbol{S}_4\|_{\mathrm{F}}^2$, the minimum Frobenius norm solution is when each $\boldsymbol{S}_i$ is all zeros.

Thus that solution has the form $\boldsymbol{G} = \boldsymbol{V} \begin{bmatrix} \boldsymbol{\Sigma}_r^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \boldsymbol{U}' = \boldsymbol{V} \boldsymbol{\Sigma}^+ \boldsymbol{U}' = \boldsymbol{A}^+$.

In words, the Moore-Penrose pseudo-inverse of $\boldsymbol{A}$ is the unique generalized inverse of $\boldsymbol{A}$ with minimal Frobenius norm. See [21] for other choices.

5.5 Procrustes analysis

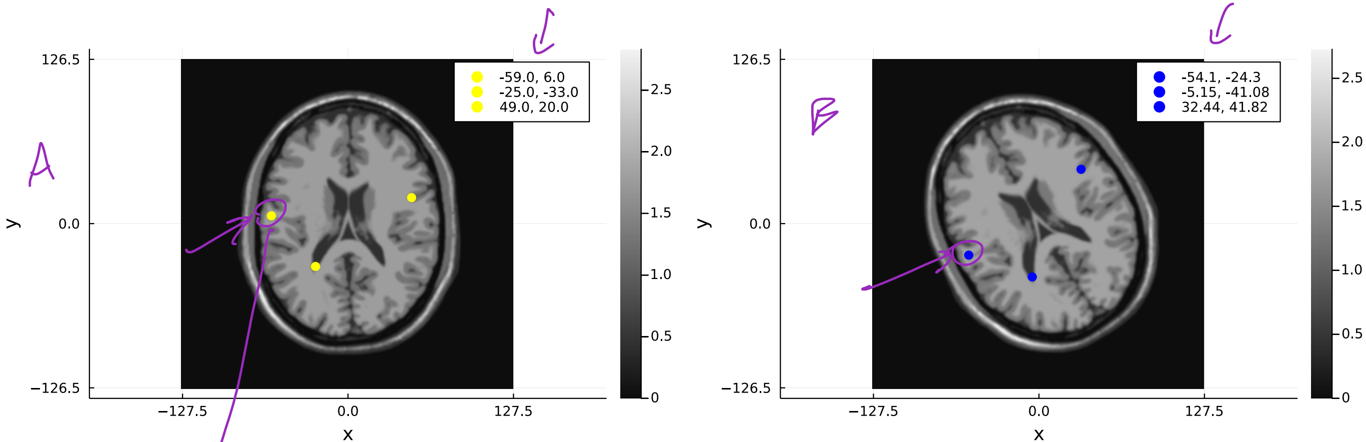
One use of matrix norms is quantifying the dissimilarity of two matrices by using a norm of their difference. This section illustrates that use by solving the **orthogonal Procrustes problem** [22, 23]. This problem provides another practical application of the **SVD** and the **Frobenius matrix norm**.

The goal of the **Procrustes problem** is to find an orthogonal matrix $\boldsymbol{Q}$ in $\mathbb{R}^{M \times M}$ that makes two other matrices $\boldsymbol{B}$ and $\boldsymbol{A}$ in $\mathbb{R}^{M \times N}$ as similar as possible by "rotating" the columns of $\boldsymbol{A}$: $A \approx QB$

$$\hat{\boldsymbol{Q}} = \underset{\boldsymbol{Q} : \boldsymbol{Q}'\boldsymbol{Q} = \boldsymbol{I}_M}{\arg\min} f(\boldsymbol{Q}), \qquad \underbrace{f(\boldsymbol{Q})}_{\hookrightarrow \text{ cost function}} \triangleq \|B - QA\|_F^2 \tag{5.25}$$

- One could use some other norm but the Frobenius is simple and natural here. (Think about why!)
- We put "rotating" in quotes because the condition $\boldsymbol{Q}'\boldsymbol{Q} = \boldsymbol{I}$ ensures that $\boldsymbol{Q}$ has orthonormal columns, but the class of matrices for which $\boldsymbol{Q}'\boldsymbol{Q} = \boldsymbol{I}$ also includes examples like $\boldsymbol{Q} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ that are not rotations.
- There are extensions that require $\det\{\boldsymbol{Q}\} = 1$ to ensure $\boldsymbol{Q}$ corresponds to a rotation [24]. ♦♦
- See p. 5.53 for several generalizations (non-square, complex, translation).

One of many motivating applications is performing **image registration** of two pictures of the same scene acquired with different sensor orientations, using a technique called **landmark registration**.



Example. Here the goal is to match (by rotation) two sets of landmark coordinates:

$$\boldsymbol{A} = \begin{bmatrix} -59 & -25 & 49 \\ 6 & -33 & 20 \end{bmatrix}, \qquad \boldsymbol{B} = \begin{bmatrix} -54.1 & -5.15 & 32.44 \\ -24.3 & -41.08 & 41.82 \end{bmatrix} \approx \boldsymbol{Q}\boldsymbol{A} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \boldsymbol{A}.$$

Here $M = 2$ and $N = 3$ and typically $M < N$ in such problems.
Here we found the landmarks manually, but there are also automatic methods [25].

To solve this problem, first analyze the cost function:

$$
\begin{aligned}
f(\boldsymbol{Q}) &= \|\boldsymbol{B}-\boldsymbol{Q}\boldsymbol{A}\|_{\mathrm{F}}^2 = \mathrm{trace}\{(\boldsymbol{B}-\boldsymbol{Q}\boldsymbol{A})'(\boldsymbol{B}-\boldsymbol{Q}\boldsymbol{A})\} && \\
&= \mathrm{trace}\{\boldsymbol{B}'\boldsymbol{B} - \boldsymbol{B}'\boldsymbol{Q}\boldsymbol{A} - \boldsymbol{A}'\boldsymbol{Q}'\boldsymbol{B} + \boldsymbol{A}'\boldsymbol{Q}'\boldsymbol{Q}\boldsymbol{A}\} && \text{expanding via } \textbf{FOIL} \\
&= \mathrm{trace}\{\boldsymbol{B}'\boldsymbol{B} - \boldsymbol{B}'\boldsymbol{Q}\boldsymbol{A} - \boldsymbol{A}'\boldsymbol{Q}'\boldsymbol{B} + \boldsymbol{A}'\boldsymbol{A}\} && \boldsymbol{Q}'\boldsymbol{Q} = \boldsymbol{I} \\
&= \mathrm{trace}\{\boldsymbol{B}'\boldsymbol{B}\} + \mathrm{trace}\{\boldsymbol{A}'\boldsymbol{A}\} - \mathrm{trace}\{\boldsymbol{A}'\boldsymbol{Q}'\boldsymbol{B}\} - \mathrm{trace}\{\boldsymbol{B}'\boldsymbol{Q}\boldsymbol{A}\} && \text{linearity} \quad tr(A+B) = tr(A) + tr(B) \\
&= \mathrm{trace}\{\boldsymbol{B}'\boldsymbol{B}\} + \mathrm{trace}\{\boldsymbol{A}'\boldsymbol{A}\} - \mathrm{trace}\{\boldsymbol{A}'\boldsymbol{Q}'\boldsymbol{B}\} - \mathrm{trace}\{(\boldsymbol{A}'\boldsymbol{Q}'\boldsymbol{B})'\} && \text{transpose} \\
&= \mathrm{trace}\{\boldsymbol{B}'\boldsymbol{B}\} + \mathrm{trace}\{\boldsymbol{A}'\boldsymbol{A}\} - 2\,\mathrm{trace}\{\boldsymbol{A}'\boldsymbol{Q}'\boldsymbol{B}\} && \text{transpose inv.} \\
&= \mathrm{trace}\{\boldsymbol{B}'\boldsymbol{B}\} + \mathrm{trace}\{\boldsymbol{A}'\boldsymbol{A}\} - 2\,\mathrm{trace}\{\boldsymbol{Q}'\boldsymbol{B}\boldsymbol{A}'\} && \text{cic} \quad \text{comm. prop. of trace.}
\end{aligned}
$$

c

So *minimizing* $f(\boldsymbol{Q})$ is equivalent to *maximizing*

$\arg\min_{Q:\, Q'Q=I} \|B-QA\|_F^2$

$$\arg\max_{Q:Q'Q=I}\; g(\boldsymbol{Q}) \triangleq \arg\max_{Q:Q'Q=I}\; \underbrace{\mathrm{trace}\{Q'BA'\}}_{g(Q)}$$

Use an **SVD** (of course!) of the $M \times M$ matrix $\boldsymbol{C} \triangleq \boldsymbol{B}\boldsymbol{A}' = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}'$ so

$$
\begin{aligned}
g(\boldsymbol{Q}) &= \mathrm{trace}\{\boldsymbol{Q}'\boldsymbol{B}\boldsymbol{A}'\} = \mathrm{trace}\{\boldsymbol{Q}'\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}'\} = \mathrm{trace}\{\boldsymbol{V}'\boldsymbol{Q}'\boldsymbol{U}\boldsymbol{\Sigma}\} \\
&= \mathrm{trace}\{\boldsymbol{W}\boldsymbol{\Sigma}\}, \qquad \boldsymbol{W} = \boldsymbol{W}(\boldsymbol{Q}) \triangleq \boldsymbol{V}'\boldsymbol{Q}'\boldsymbol{U}.
\end{aligned}
$$

Using the orthogonality of $\boldsymbol{U}$, $\boldsymbol{V}$ and $\boldsymbol{Q}$, it is clear that the $M \times M$ matrix $\boldsymbol{W}$ is orthogonal: (HW)

$$\boldsymbol{W}'\boldsymbol{W} = \boldsymbol{U}'\boldsymbol{Q}\boldsymbol{V}\boldsymbol{V}'\boldsymbol{Q}'\boldsymbol{U} = \boldsymbol{U}'\boldsymbol{Q}\boldsymbol{I}_M\boldsymbol{Q}'\boldsymbol{U} = \boldsymbol{U}'\boldsymbol{U} = \boldsymbol{I}_M.$$

unitary

We must maximize trace$\{\boldsymbol{W\Sigma}\}$ over $\boldsymbol{Q}$ orthogonal, where $\boldsymbol{W}$ depends on $\boldsymbol{Q}$ but $\boldsymbol{\Sigma}$ does not. Observe:

$w_{mm} \le 1$

$$[\boldsymbol{W\Sigma}]_{mm} = w_{mm}\sigma_m \Longrightarrow \operatorname{trace}\{\boldsymbol{W\Sigma}\} = \sum_{m=1}^{M} w_{mm}\sigma_m.$$

To proceed, we look for an upper bound for this sum. Because $\boldsymbol{W}$ is an orthogonal matrix, each of its columns have unit norm, *i.e.*, $\sum_{m=1}^{N} |w_{mn}|^2 = 1$ for all m, so $w_{mn} \le 1$ for all m, n. This inequality yields the following upper bound:

$$\operatorname{trace}\{\boldsymbol{W\Sigma}\} \le \sum_{m=1}^{M} \sigma_m = \operatorname{trace}\{\boldsymbol{I\Sigma}\}.$$

This upper bound is achieved when $\boldsymbol{W} = \boldsymbol{I}$. Now solve for $\boldsymbol{Q}$:

$$\boldsymbol{W} = \boldsymbol{V}'\boldsymbol{Q}'\boldsymbol{U} = \boldsymbol{I} \Longrightarrow \boldsymbol{V}\boldsymbol{V}'\boldsymbol{Q}'\boldsymbol{U}\boldsymbol{U}' = \boldsymbol{V}\boldsymbol{U}' \Longrightarrow \boldsymbol{Q}' = \boldsymbol{V}\boldsymbol{U}' \Longrightarrow \boxed{\hat{\boldsymbol{Q}} = \boldsymbol{U}\boldsymbol{V}'.}$$

In summary, the solution to the **orthogonal Procrustes problem** is:

$$\hat{\boldsymbol{Q}} = \mathop{\arg\min}_{\boldsymbol{Q}\,:\,\boldsymbol{Q}'\boldsymbol{Q}=\boldsymbol{I}} \|\boldsymbol{B} - \boldsymbol{Q}\boldsymbol{A}\|_{\mathrm{F}}^2 = \boldsymbol{U}\boldsymbol{V}', \text{ where } \boldsymbol{C} = \boldsymbol{B}\boldsymbol{A}' = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}'. \tag{5.26}$$

An exercise (HW) expresses $\boldsymbol{C} = \boldsymbol{Q}\boldsymbol{P}$ where $\boldsymbol{P}$ is **positive semidefinite**, using a **polar decomposition** or **polar factorization** of the square matrix $\boldsymbol{B}\boldsymbol{A}'$ [1, p. 41].

$\underset{Q:\, Q'Q=I}{\arg\min} \|B - QA\|_F^2$

$A = 0$

11. The solution to the **orthogonal Procrustes problem** is unique. (?)

A: True B: False ??

??

$C = BA' = U\Sigma V' = \sum_i \sigma_i u_i v_i'$ If BA' is full rank, is $\hat{Q}$ unique? → True

→ special case: $\Sigma = I$ → $\underbrace{(UZ)}_{\tilde{U}}\Sigma\underbrace{(VZ)'}_{\tilde{V}}$, $\Sigma = I = ZZ'$

If BA' is full rank → $\hat{Q}$ is unique

See [26] for proof.

$\tilde{U}\tilde{V}' = (UZ)(VZ)'$

If in fact $B = \tilde{Q}A$ for some unitary matrix $\tilde{Q}$, then one can verify (*cf.* next page) that $\hat{Q} = \tilde{Q}$ and that $\left\|B - \tilde{Q}A\right\|_F = 0$.

Weighted norms (Read)

Two alternatives to (5.25) involve weighted norms with positive semidefinite weighting matrices [27]:

$$\hat{Q} = \underset{Q:\, Q'Q=I}{\arg\min} f_i(Q)$$

$$f_1(Q) \triangleq \operatorname{trace}\{(B - QA) D_1 (B - QA)'\}$$

$$f_2(Q) \triangleq \operatorname{trace}\{(B - QA)' D_2 (B - QA)\}.$$

Solving f_1 is easy but f_2 requires iterative methods [27]. ♦♦

Assume $A = \frac{1}{\alpha} \tilde{Q}' B$

Sanity check (self consistency and scale invariance)

Suppose $\boldsymbol{B}$ is exactly a rotated version of the columns of $\boldsymbol{A}$, along with an additional scale factor *i.e.*, $\boldsymbol{B} = \alpha \tilde{\boldsymbol{Q}} \boldsymbol{A}$ for some orthogonal matrix $\tilde{\boldsymbol{Q}}$; equivalently $\boldsymbol{A} = \frac{1}{\alpha} \tilde{\boldsymbol{Q}}' \boldsymbol{B}$. We now verify that the Procrustes method finds the correct rotation, *i.e.*, $\hat{\boldsymbol{Q}} = \tilde{\boldsymbol{Q}}$.

Let $\boldsymbol{B} = \tilde{\boldsymbol{U}} \tilde{\boldsymbol{\Sigma}} \tilde{\boldsymbol{V}}'$ denote an SVD of $\boldsymbol{B}$. Then an SVD of $\boldsymbol{C}$ is evident by inspection:

$B = \tilde{U} \tilde{\Sigma} \tilde{V}'$

$$\boldsymbol{C} = \boldsymbol{B}\boldsymbol{A}' = \frac{1}{\alpha} \boldsymbol{B}\boldsymbol{B}' \tilde{\boldsymbol{Q}} = \frac{1}{\alpha} \underbrace{\tilde{\boldsymbol{U}} \tilde{\boldsymbol{\Sigma}} \tilde{\boldsymbol{V}}'}_{\boldsymbol{B}} \underbrace{\tilde{\boldsymbol{V}} \tilde{\boldsymbol{\Sigma}}' \tilde{\boldsymbol{U}}'}_{\boldsymbol{B}'} \tilde{\boldsymbol{Q}} = \underbrace{\tilde{\boldsymbol{U}}}_{\boldsymbol{U}} \underbrace{\frac{1}{\alpha} \tilde{\boldsymbol{\Sigma}} \tilde{\boldsymbol{\Sigma}}'}_{\boldsymbol{\Sigma}} \underbrace{\tilde{\boldsymbol{U}}' \tilde{\boldsymbol{Q}}}_{\boldsymbol{V}'}.$$

I

The Procrustes solution is indeed correct (self consistent), and **invariant** to the scale parameter α:

$$\hat{\boldsymbol{Q}} = \boldsymbol{U}\boldsymbol{V}' = (\tilde{\boldsymbol{U}})(\tilde{\boldsymbol{U}}' \tilde{\boldsymbol{Q}}) = \tilde{\boldsymbol{Q}}.$$

After finding $\hat{\boldsymbol{Q}}$, if we also want to estimate the scale, then we can solve a **linear least-squares** problem:

M

$$\arg\min_{\alpha} \left\| \boldsymbol{B} - \alpha \hat{\boldsymbol{Q}} \boldsymbol{A} \right\|_{\mathrm{F}} = \frac{\operatorname{trace}\left\{ \boldsymbol{B}\boldsymbol{A}'\hat{\boldsymbol{Q}}' \right\}}{\operatorname{trace}\{\boldsymbol{A}\boldsymbol{A}'\}} = \frac{\operatorname{trace}\{\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}'\boldsymbol{V}\boldsymbol{U}'\}}{\operatorname{trace}\{\boldsymbol{A}\boldsymbol{A}'\}} = \frac{\sum_{k=1}^{r} \sigma_k}{\|\boldsymbol{A}\|_{\mathrm{F}}^2}, \tag{5.27}$$

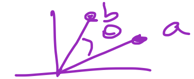
where $\{\sigma_k\}$ are the singular values of $\boldsymbol{C} = \boldsymbol{B}\boldsymbol{A}'$

$\|A\|_F^2 = \|\mathrm{vec}(A)\|_2^2$

$= \sum_{i,j} a_{ij}^2$

A HW problem will explore a real-world image registration example.
The next page provides a small concrete example.

$\arg\min_{\alpha} \sum_i \sum_j (b_{ij} - \alpha M_{ij})^2$



Example. For determining 2D image rotation, even a single nonzero point in each image suffices! (Read)
For example, suppose the first point is at $(1, 0)$ and the second point is at (x, y) where $x = 5\cos\phi$ and $y = 5\sin\phi$. (This example includes scaling by a factor of 5 just to illustrate the generality.)

Then $\boldsymbol{A} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\boldsymbol{B} = \begin{bmatrix} x \\ y \end{bmatrix}$ so

$$\boldsymbol{C} = \boldsymbol{B}\boldsymbol{A}' = \begin{bmatrix} 5\cos\phi \\ 5\sin\phi \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} \cos\phi & -q_1 \sin\phi \\ \sin\phi & q_1\cos\phi \end{bmatrix}}_{\boldsymbol{U}} \underbrace{\begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix}}_{\boldsymbol{\Sigma}} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & q_2 \end{bmatrix}}_{\boldsymbol{V}'}, \quad q_1, q_2 \in \{\pm 1\}.$$

Here $\boldsymbol{C}$ is a simple outer product so finding a (full!) SVD by hand was easy.
In fact we found four SVDs, corresponding to different signs for $\boldsymbol{u}_2$ and $\boldsymbol{v}_2$.
For each of these SVDs, the optimal rotation matrix per (5.26) is

$$\hat{\boldsymbol{Q}} = \boldsymbol{U}\boldsymbol{V}' = \begin{bmatrix} \cos\phi & -q_1 \sin\phi \\ \sin\phi & q_1\cos\phi \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & q_2 \end{bmatrix} = \begin{bmatrix} \cos\phi & -q\sin\phi \\ \sin\phi & q\cos\phi \end{bmatrix},$$

where $q \triangleq q_1 q_2 \in \{\pm 1\}$. The two Procrustes solutions here (for $q = \pm 1$) both have the correct $\cos\phi$ in the upper left and both exactly satisfy $\boldsymbol{B} = 5\boldsymbol{Q}\boldsymbol{A}$.
So there are two Procrustes solutions that fit the data exactly, one of which (for $q = 1$) corresponds to a rotation matrix, and the other of which (for $q = -1$) has sign flip for the second coordinate.
In 2D, any rotation matrix is a unitary matrix, but the converse is not true!

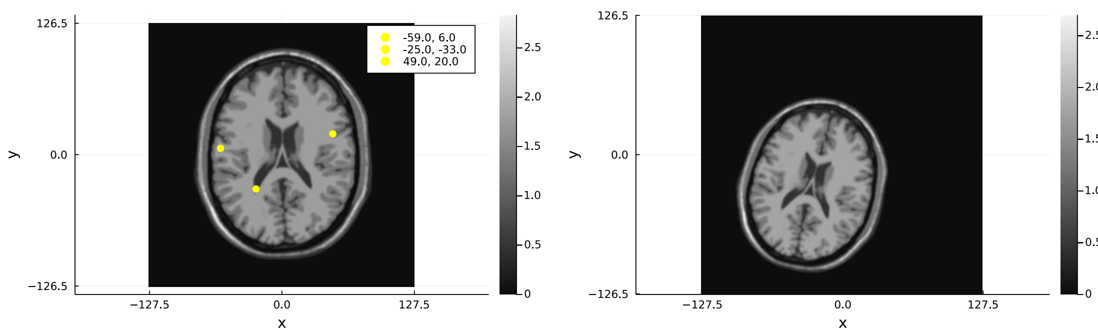
Exercise. Explore what happens with two points that are: colinear, symmetric around zero, or non-colinear.

Procrustes Generalizations: non-square, complex, with translation (Read)

This section generalizes the Procrustes problem (5.25) in three ways: we consider complex data, we account for a possible translation, and we allow $\boldsymbol{Q}$ to be non-square, meaning that $\boldsymbol{B}$ and $\boldsymbol{A}$ can have different numbers of rows.

Procrustes was a figure in Greek mythology who "stretched" people to fit an iron bed. An even more general version of the Procrustes problem considers a scaling factor (stretching). (HW) (See [wiki] figure.)

A practical application of such generalizations are in landmark-based image registration where one can encounter rotation, translation, and scale factors (due to different pixel sizes).



Here we assume $\boldsymbol{B} \in \mathbb{F}^{M \times N}$ but $\boldsymbol{A} \in \mathbb{F}^{K \times N}$ so $\boldsymbol{Q} \in \mathbb{F}^{M \times K}$.

We still want $\boldsymbol{Q}$ to have orthonormal columns, so we must have $1 \leq K \leq M$.

Define. The **Stiefel manifold** $\mathcal{V}_K(\mathbb{F}^M)$ is the set of $M \times K$ matrices having orthonormal columns

$$\mathcal{V}_K(\mathbb{F}^M) = \left\{ \boldsymbol{Q} \in \mathbb{F}^{M \times K} \;:\; \boldsymbol{Q}'\boldsymbol{Q} = \boldsymbol{I}_K \right\}.$$

set! not subspace!

$$\underbrace{\begin{bmatrix} q_1' \\ \vdots \\ q_K' \end{bmatrix}}_{Q'} \underbrace{\begin{bmatrix} q_1 & \cdots & q_K \end{bmatrix}}_{Q} = \begin{pmatrix} q_1'q_1 & & 0 \\ & \ddots & \\ 0 & & q_K'q_K \end{pmatrix}$$

Special cases:
$\mathcal{V}_M(\mathbb{R}^M)$ is the set of $M \times M$ orthogonal matrices
$\mathcal{V}_M(\mathbb{C}^M)$ is the set of $M \times M$ unitary matrices

If $\boldsymbol{Q} \in \mathcal{V}_K(\mathbb{C}^M)$ then $\boldsymbol{Q}$ is the first K columns of some $M \times M$ unitary matrix.

In many practical applications of the **Procrustes problem**, there can be both rotation and an unknown **translation** between the two sets of coordinates. Instead of the model $\boldsymbol{B}_{:,n} \approx \boldsymbol{Q}\boldsymbol{A}_{:,n}$ a more realistic model is $\boldsymbol{B}_{:,n} \approx \boldsymbol{Q}\boldsymbol{A}_{:,n} + \boldsymbol{d}$ where $\boldsymbol{d} \in \mathbb{F}^M$ is an unknown **displacement vector**. In matrix form:

$$\boldsymbol{B} \approx \boldsymbol{Q}\boldsymbol{A} + \boldsymbol{d}\boldsymbol{1}_N'.$$

Now we must determine both a matrix $\boldsymbol{Q} \in \mathcal{V}_K(\mathbb{F}^M)$ in the Stiefel manifold, and the vector $\boldsymbol{d} \in \mathbb{C}^M$ by a double minimization using a Frobenius norm:

$$(\hat{\boldsymbol{Q}}, \hat{\boldsymbol{d}}) \triangleq \underset{\boldsymbol{Q} \in \mathcal{V}_K(\mathbb{F}^M)}{\arg\min} \; \underset{\boldsymbol{d} \in \mathbb{F}^M}{\arg\min} \, g(\boldsymbol{d}, \boldsymbol{Q}), \qquad g(\boldsymbol{d}, \boldsymbol{Q}) \triangleq \left\| \boldsymbol{B} - (\boldsymbol{Q}\boldsymbol{A} + \boldsymbol{d}\boldsymbol{1}_N') \right\|_{\mathrm{F}}^2.$$

We first focus on the inner minimization over the displacement $\boldsymbol{d}$ for any given $\boldsymbol{Q}$:

$$\begin{aligned} g(\boldsymbol{d}, \boldsymbol{Q}) &= \|\boldsymbol{B} - (\boldsymbol{Q}\boldsymbol{A} + \boldsymbol{d}\mathbf{1}_N')\|_{\mathrm{F}}^2 = \mathsf{trace}\{(\boldsymbol{Z} - \boldsymbol{d}\mathbf{1}_N')'(\boldsymbol{Z} - \boldsymbol{d}\mathbf{1}_N')\}, \quad \boldsymbol{Z} \triangleq \boldsymbol{B} - \boldsymbol{Q}\boldsymbol{A} \\ &= \mathsf{trace}\{\boldsymbol{Z}'\boldsymbol{Z}\} - \mathsf{trace}\{\boldsymbol{Z}'\boldsymbol{d}\mathbf{1}_N'\} - \mathsf{trace}\{\mathbf{1}_N\boldsymbol{d}'\boldsymbol{Z}\} + \mathsf{trace}\{\mathbf{1}_N\boldsymbol{d}'\boldsymbol{d}\mathbf{1}_N'\} \\ &= \mathsf{trace}\{\boldsymbol{Z}'\boldsymbol{Z}\} - \mathsf{trace}\{\mathbf{1}_N'\boldsymbol{Z}'\boldsymbol{d}\} - \mathsf{trace}\{\boldsymbol{d}'\boldsymbol{Z}\mathbf{1}_N\} + \mathsf{trace}\{\boldsymbol{d}'\boldsymbol{d}\mathbf{1}_N'\mathbf{1}_N\} \\ &= \mathsf{trace}\{\boldsymbol{Z}'\boldsymbol{Z}\} - \mathbf{1}_N'\boldsymbol{Z}'\boldsymbol{d} - \boldsymbol{d}'\boldsymbol{Z}\mathbf{1}_N + N\boldsymbol{d}'\boldsymbol{d} \\ &= \mathsf{trace}\{\boldsymbol{Z}'\boldsymbol{Z}\} - \frac{1}{N}\|\boldsymbol{Z}\mathbf{1}_N\|_2^2 + N\left\|\boldsymbol{d} - \frac{1}{N}\boldsymbol{Z}\mathbf{1}_N\right\|_2^2. \end{aligned}$$

It is clear from this expression that the optimal estimate of the displacement $\boldsymbol{d}$ for any $\boldsymbol{Q}$ is:

$$\hat{\boldsymbol{d}}(\boldsymbol{Q}) = \frac{1}{N}\boldsymbol{Z}\mathbf{1}_N = \frac{1}{N}(\boldsymbol{B} - \boldsymbol{Q}\boldsymbol{A})\mathbf{1}_N.$$

Now to find the optimal matrix $\boldsymbol{Q}$ we must solve the outer minimization:

$$\hat{\boldsymbol{Q}} \triangleq \underset{\boldsymbol{Q} \in \mathcal{V}_K(\mathbb{F}^M)}{\arg\min} \, f(\boldsymbol{Q}), \qquad f(\boldsymbol{Q}) \triangleq g\left(\hat{\boldsymbol{d}}(\boldsymbol{Q}), \boldsymbol{Q}\right)$$

$$\begin{aligned}
f(\boldsymbol{Q}) &= \mathsf{trace}\{(\boldsymbol{B} - \boldsymbol{Q}\boldsymbol{A})'(\boldsymbol{B} - \boldsymbol{A}\boldsymbol{Q})\} - \frac{1}{N} \|(\boldsymbol{B} - \boldsymbol{Q}\boldsymbol{A})\mathbf{1}_N\|_2^2 \\
&\stackrel{c}{=} -2\,\mathrm{real}\{\mathsf{trace}\{\boldsymbol{Q}'\boldsymbol{B}\boldsymbol{A}'\}\} + \frac{2}{N}\,\mathrm{real}\{\mathbf{1}_N'\boldsymbol{A}'\boldsymbol{Q}'\boldsymbol{B}\mathbf{1}_N\} \\
&= -2\,\mathrm{real}\{\mathsf{trace}\{\boldsymbol{Q}'\boldsymbol{B}\boldsymbol{A}'\}\} + 2\,\mathrm{real}\left\{\mathsf{trace}\left\{\boldsymbol{Q}'\boldsymbol{B}\frac{1}{N}\mathbf{1}_N\mathbf{1}_N'\boldsymbol{A}'\right\}\right\} \\
&= -2\,\mathrm{real}\left\{\mathsf{trace}\left\{\boldsymbol{Q}'\tilde{\boldsymbol{C}}\right\}\right\}, \quad \tilde{\boldsymbol{C}} \triangleq \underbrace{\boldsymbol{B}\boldsymbol{M}\boldsymbol{A}'}_{M \times K}, \quad \boldsymbol{M} \triangleq \boldsymbol{I} - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N',
\end{aligned}$$

where $\stackrel{c}{=}$ means "equal to within constant terms that are irrelevant for minimization."
After finding a (full) **SVD** $\tilde{\boldsymbol{C}} = \underbrace{\boldsymbol{U}}_{M \times M}\underbrace{\boldsymbol{\Sigma}}_{M \times K}\underbrace{\boldsymbol{V}'}_{K \times K}$, we want:

$$\hat{\boldsymbol{Q}} = \underset{\boldsymbol{Q} \in \mathcal{V}_K(\mathbb{F}^M)}{\arg\max} \, \mathrm{real}\left\{\mathsf{trace}\left\{\boldsymbol{Q}'\tilde{\boldsymbol{C}}\right\}\right\}$$

where using the Frobenius inner product inequality (5.12):

$$\begin{aligned}
\mathrm{real}\left\{\mathsf{trace}\left\{\boldsymbol{Q}'\tilde{\boldsymbol{C}}\right\}\right\} &= \mathrm{real}\{\mathsf{trace}\{\boldsymbol{Q}'\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}'\}\} = \mathrm{real}\{\mathsf{trace}\{\boldsymbol{W}'\boldsymbol{\Sigma}\}\}, \quad \text{where } \boldsymbol{W} \triangleq \boldsymbol{U}'\boldsymbol{Q}\boldsymbol{V} \in \mathcal{V}_K(\mathbb{F}^M) \\
&= \mathrm{real}\{\langle \boldsymbol{\Sigma}, \boldsymbol{W} \rangle\} \le |\langle \boldsymbol{\Sigma}, \boldsymbol{W} \rangle| \le \|\mathsf{vec}(\boldsymbol{\Sigma})\|_1 \|\mathsf{vec}(\boldsymbol{W})\|_\infty = \|\boldsymbol{\Sigma}\|_* \|\mathsf{vec}(\boldsymbol{W})\|_\infty .
\end{aligned}$$

Because $\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_K \\ \mathbf{0}_{(M-K)\times K} \end{bmatrix}$ is rectangular diagonal, the matrix $\boldsymbol{W} = \begin{bmatrix} \boldsymbol{I}_K \\ \mathbf{0}_{(M-K)\times K} \end{bmatrix} \in \mathcal{V}_K(\mathbb{F}^M)$ achieves the upper bound. Solving for $\boldsymbol{Q}$ yields

$$\hat{\boldsymbol{Q}} = \boldsymbol{U}\boldsymbol{W}\boldsymbol{V}' = \boldsymbol{U} \begin{bmatrix} \boldsymbol{I}_K \\ \mathbf{0}_{(M-K)\times K} \end{bmatrix} \boldsymbol{V}' = \boldsymbol{U}_K \boldsymbol{V}',$$

where $\boldsymbol{U}_K$ denotes the first K columns of the $M \times M$ matrix $\boldsymbol{U}$.

In summary, the optimal $\boldsymbol{Q}$ is

$$\boldsymbol{Q} = \boldsymbol{U}_K \boldsymbol{V}', \text{ where } \tilde{\boldsymbol{C}} \triangleq \boldsymbol{B}\boldsymbol{M}\boldsymbol{A}' = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}'.$$

The matrix $\boldsymbol{M}$ is called a "de-meaning" or "centering" operator because $\boldsymbol{y} = \boldsymbol{M}\boldsymbol{x}$ subtracts the mean of $\boldsymbol{x}$ from each element of $\boldsymbol{x}$. In code: `y = x .- mean(x)`

The de-meaning matrix $\boldsymbol{M}$ is a symmetric **idempotent matrix** so $\boldsymbol{M} = \boldsymbol{M}\boldsymbol{M}'$ and we can rewrite $\tilde{\boldsymbol{C}}$ above as $\tilde{\boldsymbol{C}} = (\boldsymbol{B}\boldsymbol{M})(\boldsymbol{A}\boldsymbol{M})' = \tilde{\boldsymbol{B}}\tilde{\boldsymbol{A}}'$ where $\tilde{\boldsymbol{A}} \triangleq \boldsymbol{A}\boldsymbol{M}$, $\tilde{\boldsymbol{B}} \triangleq \boldsymbol{B}\boldsymbol{M}$ are versions of $\boldsymbol{A}$ and $\boldsymbol{B}$ where each column has its mean subtracted out.

In words, to find the optimal rotation matrix when there is possible translation, we first de-mean each column of $\boldsymbol{A}$ and $\boldsymbol{B}$, and then compute the usual SVD of $\tilde{\boldsymbol{B}}\tilde{\boldsymbol{A}}'$ and use the left and right bases via $\boldsymbol{Q} = \boldsymbol{U}_K \boldsymbol{V}'$.

Exercise. Do a sanity check in the case where $\boldsymbol{B} = \alpha \boldsymbol{Q}\boldsymbol{A} + \boldsymbol{d}\mathbf{1}_N'$.

Subspace / span comparisons (Read)

Another application of the **orthogonal Procrustes problem** is quantifying the "alignment" between two subspace bases.

Suppose $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ are $M \times N$ matrices that we think span the same (or similar) subspace in $\mathbb{F}^M$. In general it does not make sense to use $d(\boldsymbol{B}_1, \boldsymbol{B}_2) = \|\boldsymbol{B}_1 - \boldsymbol{B}_2\|_{\mathrm{F}}$ as a measure if dissimilarity because we could have $\mathcal{R}(\boldsymbol{B}_1) = \mathcal{R}(\boldsymbol{B}_2)$ even if $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ are themselves different, *e.g.*, if $\boldsymbol{B}_1 = -\boldsymbol{B}_2$.

A more useful measure of dissimilarity involves first rotating the basis for one subspace to be as similar to the other as possible, and then examining the difference, *i.e.*:

$$d(\boldsymbol{B}_1, \boldsymbol{B}_2) \triangleq \min_{\boldsymbol{Q} \in \mathcal{V}_N(\mathbb{F}^N)} \|\boldsymbol{B}_1 - \boldsymbol{B}_2 \boldsymbol{Q}'\|_{\mathrm{F}} = \min_{\boldsymbol{Q} \in \mathcal{V}_N(\mathbb{F}^N)} \|\boldsymbol{B}_1' - \boldsymbol{Q} \boldsymbol{B}_2'\|_{\mathrm{F}} .$$

The best $\boldsymbol{Q}$ is $\hat{\boldsymbol{Q}} = \boldsymbol{U}\boldsymbol{V}'$ where $\boldsymbol{C} = \boldsymbol{B}_1' \boldsymbol{B}_2 = \boldsymbol{U} \boldsymbol{\Sigma} \boldsymbol{V}'$, so the simplified dissimilarity measure is

$$d(\boldsymbol{B}_1, \boldsymbol{B}_2) = \|\boldsymbol{B}_1 - \boldsymbol{B}_2 \boldsymbol{V} \boldsymbol{U}'\|_{\mathrm{F}} .$$

If the $\boldsymbol{B}$ matrices are not in the **Stiefel manifold**, then one should include a scale factor like (5.27).

Practical implementation

The solution to the Procrustes problem requires just a couple JULIA statements. The key ingredient is simply the `svd` command. See the example notebook:

https://web.eecs.umich.edu/~fessler/course/551/julia/demo/05_procrustes1.html
https://web.eecs.umich.edu/~fessler/course/551/julia/demo/05_procrustes1.ipynb

5.6 Summary

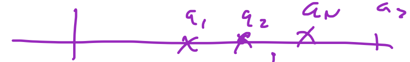
- We use **vector norms** and **matrix norms** are used to measure sizes and distances.
- Some **matrix norms** are essentially just vector norms in terms of $\text{vec}(\boldsymbol{A})$, some matrix norms satisfy the important **sub-multiplicative** property, and **operator norms** are induced by vector norms.
- Many of the matrix norms can be expressed in terms of singular values, and those are **unitarily invariant**.
- Classical methods (like linear LS) use 2-norms, but many modern methods use other norms. One vector norm of recent interest is the **ordered weighted ℓ_1 (OWL)** norm [28].
- We assess **convergence** of a sequence of vectors or matrices using norms.
- The **spectral radius** is a related quantity for square matrices, where $\sigma_1(\boldsymbol{A}) = \sqrt{\sigma_1(\boldsymbol{A}'\boldsymbol{A})} = \sqrt{\rho(\boldsymbol{A}'\boldsymbol{A})}$.
- The **Moore-Penrose pseudo-inverse** is the **generalized inverse** having minimum **Frobenius norm**.

The **orthogonal Procrustes problem** has an **SVD**-based solution:

$$\hat{\boldsymbol{Q}} = \underset{\boldsymbol{Q}\,:\,\boldsymbol{Q}'\boldsymbol{Q}=\boldsymbol{I}_M}{\arg\min} \, \|\boldsymbol{B} - \boldsymbol{Q}\boldsymbol{A}\|_{\mathrm{F}}^2 = \boldsymbol{U}\boldsymbol{V}', \qquad \boldsymbol{C} = \boldsymbol{B}\boldsymbol{A}' = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}'.$$

- The solution is invariant to scaling factors $\alpha\boldsymbol{Q}\boldsymbol{A}$
- Unknown displacement (translation) simply requires de-meaning $\boldsymbol{A}$ and $\boldsymbol{B}$ before doing SVD
- Displacement estimate (if needed) is $\frac{1}{N}\left(\boldsymbol{B} - \hat{\boldsymbol{Q}}\boldsymbol{A}\right)\mathbf{1}_N$.

Deriving the solution to this problem used *many* of the tools discussed so far: Frobenius norm, matrix trace and its properties, SVD, matrix/vector algebra.



Exercise. Suppose $\boldsymbol{A}$ and $\boldsymbol{B}$ are both real $1 \times N$ vectors (each with mean 0 for simplicity).
How can we interpret the orthogonal Procrustes solution in this case geometrically?
Hint: What is SVD of $\boldsymbol{B}\boldsymbol{A}'$ here? If $\boldsymbol{A} = \boldsymbol{x}'$ and $\boldsymbol{B} = \boldsymbol{y}'$ where $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$, then $\boldsymbol{B}\boldsymbol{A}' = \boldsymbol{y}'\boldsymbol{x} =$ scalar $\underbrace{\operatorname{sgn}(\boldsymbol{y}'\boldsymbol{x})}_{\boldsymbol{U}} \underbrace{|\boldsymbol{y}'\boldsymbol{x}|}_{\sigma_1} \underbrace{1}_{\boldsymbol{V}}$
so $\boldsymbol{Q} = \boldsymbol{U}\boldsymbol{V}' = \operatorname{sgn}(\boldsymbol{y}'\boldsymbol{x}) = \pm 1$. Here the "rotation" is just possibly negating the sign to match in 1D.

Exercise. What if $\boldsymbol{B} = \mathrm{e}^{\imath\phi} \boldsymbol{A}$? (in class if possible)

(1) Form $C = BA' = e^{i\phi} AA'$
(2) Take SVD $C = U\Sigma V' = e^{i\phi} VV'$
(3) Find $\hat{Q}$ $\hat{Q} = UV' = e^{i\phi} VV' = e^{i\phi} I$

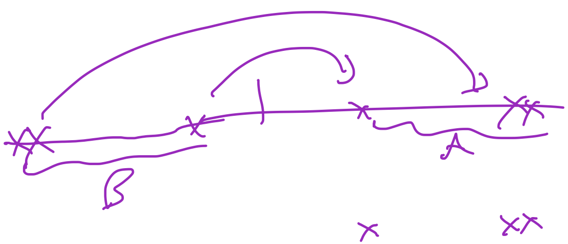
$\boxed{??}$

Challenge. (Return to this after studying Ch. 8.) The Frobenius norm is not robust to **outlier** data. Develop a formulation using something like an ℓ_1 norm instead, to provide better robustness [29].

$A = [1 \quad 5 \quad 5]$

$B = [-1 \quad -5 \quad -5]$

$\|B - QA\|_F^2$



Bibliography

[1] A. J. Laub. *Matrix analysis for scientists and engineers*. Soc. Indust. Appl. Math., 2005 (cit. on pp. 5.2, 5.3, 5.19, 5.26, 5.32, 5.49).

[2] D. G. Luenberger. *Optimization by vector space methods*. New York: Wiley, 1969 (cit. on p. 5.4).

[3] C. Liu, Z. Li, J. Yuan, and Y. Xu. *NeLF: practical novel view synthesis with neural light field*. 2021 (cit. on p. 5.10).

[4] A. Bjorck and G. H. Golub. "Numerical methods for computing angles between linear subspaces". In: *Mathematics of Computation* 27.123 (July 1973), 579–94 (cit. on p. 5.16).

[5] W. Leeb and E. Romanov. *Optimal spectral shrinkage and PCA with heteroscedastic noise*. 2019 (cit. on p. 5.16).

[6] V. Eijkhout and P. Vassilevski. "The role of the strengthened Cauchy-Buniakowskii-Schwarz inequality in multilevel methods". In: *SIAM Review* 33.3 (Sept. 1991), 405–19 (cit. on p. 5.16).

[7] L. Mirsky. "On the trace of matrix products". In: *Mathematische Nachrichten* 20.3-6 (1959), 171–4 (cit. on p. 5.18).

[8] R. H. Chan and M. K. Ng. "Conjugate gradient methods for Toeplitz systems". In: *SIAM Review* 38.3 (Sept. 1996), 427–82 (cit. on p. 5.21).

[9] V-S. Chellaboina and W. M. Haddad. "Is the Frobenius matrix norm induced?" In: *IEEE Trans. Auto. Control* 40.12 (Dec. 1995), 2137–9 (cit. on p. 5.21).

[10] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge: Cambridge Univ. Press, 1985 (cit. on p. 5.21).

[11] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado. "Sparse solutions to linear inverse problems with multiple measurement vectors". In: *IEEE Trans. Sig. Proc.* 53.7 (July 2005), 2477–88 (cit. on p. 5.22).

[12] J. A. Tropp, A. C. Gilbert, and M. J. Strauss. "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit". In: *Signal Processing* 86.3 (Mar. 2006), 572–88 (cit. on p. 5.22).

[13] S. Feizi and D. Tse. *Maximally correlated principal component analysis*. 2017 (cit. on p. 5.27).

[14] D. Cullina, P. Mittal, and N. Kiyavash. "Fundamental limits of database alignment". In: *Intl. Symp. on Information Theory*. 2018, 651–5 (cit. on p. 5.27).

[15] C. Xu, B. Yang, F. Guo, W. Zheng, and P. Poignet. "Sparse-view CBCT reconstruction via weighted Schatten p-norm minimization". In: *Optics Express* 28.24 (Nov. 2020), 35469–82 (cit. on p. 5.27).

[16] R. Bhatia. *Matrix analysis*. New York: Springer, 1997 (cit. on pp. 5.28, 5.35).

[17] K. Fan. "Maximum properties and inequalities for the eigenvalues of completely continuous operators". In: *Proc. Natl. Acad. Sci.* 37.11 (Nov. 1951), 760–6 (cit. on p. 5.33).

[18] R. Bhatia and F. Kittaneh. "Notes on matrix arithmetic-geometric mean inequalities". In: *Linear Algebra and its Applications* 308.1-3 (Mar. 2000), 203–11 (cit. on p. 5.36).

[19] R. Pascanu, T. Mikolov, and Y. Bengio. "On the difficulty of training recurrent neural networks". In: *pmlr* 28.3 (2013), 1310–8 (cit. on p. 5.38).

[20] D. Gilton, G. Ongie, and R. Willett. "Neumann networks for inverse problems in imaging". In: *IEEE Trans. Computational Imaging* 6 (2020), 328–43 (cit. on p. 5.43).

[21] I. Dokmanic and Remi Gribonval. *Beyond Moore-Penrose part I: Generalized inverses that minimize matrix norms*. 2017 (cit. on p. 5.45).

[22] P. H. Schonemann. "A generalized solution of the orthogonal Procrustes problem". In: *Psychometrika* 31.1 (Mar. 1966), 1–10 (cit. on p. 5.46).

[23] J. C. Gower and G. B. Dijksterhuis. *Procrustes problems*. Oxford, 2004 (cit. on p. 5.46).

[24] S. Ahmed and I. M. Jaimoukha. "A relaxation-based approach for the orthogonal Procrustes problem with data uncertainties". In: *Proc. UKACC Intl. Conf. Control*. 2012, 906–11 (cit. on p. 5.46).

[25] F. C. Ghesu, B. Georgescu, S. Grbic, A. K. Maier, J. Hornegger, and D. Comaniciu. "Robust multi-scale anatomical landmark detection in incomplete 3D-CT data". In: *Medical Image Computing and Computer-Assisted Intervention*. 2017, 194–202 (cit. on p. 5.47).

[26] S. Ravishankar and Y. Bresler. "l_0 sparsifying transform learning with efficient optimal updates and convergence guarantees". In: *IEEE Trans. Sig. Proc.* 63.9 (May 2015), 2389–404 (cit. on p. 5.50).

[27] M. A. Koschat and D. F. Swayne. "A weighted Procrustes criterion". In: *Psychometrika* 56.2 (June 1991), 229–39 (cit. on p. 5.50).

[28] M. A. T. Figueiredo and R. D. Nowak. "Ordered weighted l1 regularized regression with strongly correlated covariates: Theoretical aspects". In: *aistats*. 2016, 930–8 (cit. on p. 5.59).

[29] P. J. F. Groenen, P. Giaquinto, and H. A. L. Kiers. "An improved majorization algorithm for robust Procrustes analysis". In: *New Developments in Classification and Data Analysis: Proceedings of the Meeting of the Classification and Data Analysis Group (CLADAG) of the Italian Statistical Society, University of Bologna*. Berlin: Springer, 2005, pp. 151–8 (cit. on p. 5.60).