CO233 - Computational Techniques

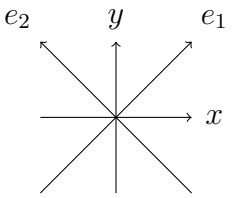
15th January 2020

Vector and Matrix Norms

An orthonormal basis of $\mathbb{R}^n$ are unit vectors that are pairwise mutually perpendicular; such that for $(e_1, \cdots, e_n)$;

- $e_i \cdot e_i = 1$
- $e_i \cdot e_j = 0$, if $i \neq j$

The standard canonical basis of $\mathbb{R}^3$ are the i, j, k vectors, and similar in $\mathbb{R}^2$. However, we can form another orthonormal basis of $\mathbb{R}^2$ by bisecting the angles as such;



If we take a vector $\boldsymbol{v} \in \mathbb{R}^n$, the Euclidean norm (or the ℓ_2-norm) is defined as such;

$$\|\boldsymbol{v}\|_2 = \sqrt{\sum_{i=1}^{n} v_i^2}$$

A norm, a mapping $\| \ \| : \mathbb{R}^n \to \mathbb{R}^+$, must satisfy these 3 axioms;

(i) $\|\boldsymbol{v}\| > 0$ given that $\boldsymbol{v} \neq \boldsymbol{0}$

(ii) $\|\lambda \boldsymbol{v}\| = |\lambda| \ \|\boldsymbol{v}\|$

(iii) $\|\boldsymbol{v} + \boldsymbol{w}\| \leq \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$ (triangular inequality)

Some other (ℓ_p) norms are defined as follows;

$$\ell_1\text{-norm } \|\boldsymbol{v}\|_1 = \sum_{i=1}^{n} |v_i|$$

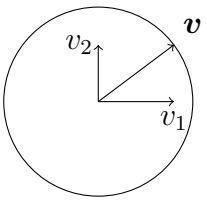
$$\ell_\infty\text{-norm } \|\boldsymbol{v}\|_\infty = \max\{|v_i| : 1 \leq i \leq n\}$$

$$\ell_p\text{-norm } \|\boldsymbol{v}\|_p = \left(\sum_{i=1}^{n} |v_i|^p \right)^{\frac{1}{p}}$$

In each dimension, we have the following;

- $n = 1$
  - $\|\boldsymbol{v}\|_1 = |v| = |v|$
  - $\|\boldsymbol{v}\|_2 = \sqrt{v^2} = |v|$
  - $\|\boldsymbol{v}\|_\infty = \max\{|v|\} = |v|$
- $n = 2$

  We can represent this geometrically as such;



In our case $\|\boldsymbol{v}\|_\infty = \max\{|v_1|, |v_2|\} = |v_1|$, but the point is that it's either of the "sides" of the triangle. Obviously, $\|\boldsymbol{v}\|_2 \geq \|\boldsymbol{v}\|_\infty$, as it's the hypotenuse of the triangle, and similarly, $\|\boldsymbol{v}\|_1 \geq \|\boldsymbol{v}\|_2$, due to the triangle inequality. Therefore we have $\|\boldsymbol{v}\|_1 \geq \|\boldsymbol{v}\|_2 \geq \|\boldsymbol{v}\|_\infty$.

Even if the orthonormal base changes, the Euclidean norm stays the same, whereas the other norms can change. As such, we can say the ℓ_2-norm is invariant under an **orthogonal transformation** (a basis change from an orthonormal bases to another orthonormal bases).

- $n = ?$ (general)

  The goal is to prove $\|\boldsymbol{v}\|_\infty \leq \|\boldsymbol{v}\|_2 \leq \|\boldsymbol{v}\|_1$. If we first take the squares of all of them, such that we have the following;

$$\|\boldsymbol{v}\|_1^2 = \left(\sum_{i=1}^{n} |v_i|\right)^2$$
$$\|\boldsymbol{v}\|_2^2 = \sum_{i=1}^{n} |v_i|^2$$
$$\|\boldsymbol{v}\|_\infty^2 = (\max\{|v_i| : 1 \leq i \leq n\})^2$$

  Since the ℓ_∞-norm corresponds to a single v_i, it's obvious that the following inequality holds (since the ℓ_2-norm squared has all the other terms squared, as well as the ℓ_∞-norm squared);

$$\|\boldsymbol{v}\|_2^2 = \sum_{i=1}^{n} |v_i|^2 \geq (\max\{|v_i| : 1 \leq i \leq n\})^2 = \|\boldsymbol{v}\|_\infty^2 \Rightarrow \|\boldsymbol{v}\|_2 \geq \|\boldsymbol{v}\|_\infty$$

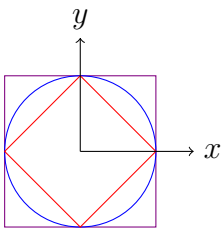
  To prove the other inequality, we see that the square of the sum of absolutes is greater than the sum of the squares, as the square of the sum contains the cross terms (which will be positive).

$$\|\boldsymbol{v}\|_2^2 = \sum_{i=1}^{n} |v_i|^2 \leq \left(\sum_{i=1}^{n} |v_i|\right)^2 = \|\boldsymbol{v}\|_\infty^1 \Rightarrow \|\boldsymbol{v}\|_2 \leq \|\boldsymbol{v}\|_1$$

  As such, we can conclude that $\|\boldsymbol{v}\|_\infty \leq \|\boldsymbol{v}\|_2 \leq \|\boldsymbol{v}\|_1$ in any dimension. ■

Tutorial Question

Find the locus of vectors such that $\|\boldsymbol{v}\|_p \leq 1$, for $p = 1, 2, \infty$ in $n = 2$;



Imagine they're all shaded from the border to the origin.

ℓ_p-norm

Our goal is to show that as $p \to \infty$, we get the definition of the ℓ_∞-norm previously stated. Take a vector $\boldsymbol{v} \in \mathbb{R}^n$, where both $\boldsymbol{v}$ and n are fixed.

$$\|\boldsymbol{v}\|_p^p = \sum_{i=1}^{n} |v_i|^p$$

Obviously, this is greater than or equal to $\|\boldsymbol{v}\|_\infty^p$, as it would only be a single $|v_i|^p$. Similarly, it must be less than or equal to $n|v_i|^p$, as v_i is the maximum of all the components.

$$\|\boldsymbol{v}\|_\infty^p \leq \|\boldsymbol{v}\|_p^p = \sum_{i=1}^{n} |v_i|^p \leq n\|\boldsymbol{v}\|_\infty^p$$

Taking everything to the power of $\frac{1}{p}$, we obtain the following result (note that $p > 0$ hence the signs don't change);

$$\|\boldsymbol{v}\|_\infty \leq \|\boldsymbol{v}\|_p =\leq n^{\frac{1}{p}}\|\boldsymbol{v}\|_\infty$$

As $p \to \infty$, since $n \geq 2$ ($n = 1$ is shown to collapse to the same component), we have $n^{\frac{1}{p}} \to 1$, which sandwiches the middle term.

Some Proposition (ℓ_∞-norm vs ℓ_2-norm)

The proposition is as follows; for a vector $\boldsymbol{v} \in \mathbb{R}^n$, $\|\boldsymbol{v}\|_2 \leq \sqrt{n}\|\boldsymbol{v}\|_\infty$. To show this, we know that each of $|v_i|$ is less than or equal to $\|\boldsymbol{v}\|_\infty$, by definition of the maximum. The same can be said for $|v_i|^2$, vs $\|\boldsymbol{v}\|_\infty^2$. Taking square roots, we have the following;

$$\|\boldsymbol{v}\|_2^2 = \sum_{i=1}^{n} |v_i|^2 \leq n\|\boldsymbol{v}\|_\infty^2 \Rightarrow \|\boldsymbol{v}\|_2 \leq \sqrt{n}\|\boldsymbol{v}\|_\infty$$

To show this holds similarly for $\|\boldsymbol{v}\|_1 \leq \sqrt{n}\|\boldsymbol{v}\|_2$, we employ the Cauchy-Schwarz inequality, which states $|\boldsymbol{x} \cdot \boldsymbol{y}| \leq \|\boldsymbol{x}\|_2\|\boldsymbol{y}\|_2$. The Cauchy-Schwarz inequality uses the fact that $\boldsymbol{x} \cdot \boldsymbol{y} = \|\boldsymbol{x}\|_2\|\boldsymbol{y}\|_2 \cos\theta$. To do this, we need to define a sign function $\text{sgn} : \mathbb{R} \to \{1, -1\}$ as follows;

$$\text{sgn } x = \begin{cases} 1 & x \geq 0 \\ -1 & x < 0 \end{cases}$$

We also need to craft a vector $\boldsymbol{w}$, as follows;

$$\begin{aligned}
w_i &= \frac{\text{sgn } v_i}{\sqrt{n}} && 1 \leq i \leq n \\
\boldsymbol{v} \cdot \boldsymbol{w} &= \sum_{i=1}^{n} v_i w_i \\
&= \sum_{i=1}^{n} \frac{v_i \cdot \text{sgn } v_i}{\sqrt{n}} && \text{product of same sign becomes positive} \\
&= \sum_{i=1}^{n} \frac{|v_i|}{\sqrt{n}} \\
&= \sqrt{n} \sum_{i=1}^{n} |v_i| \\
&= \sqrt{n}\|\boldsymbol{v}\|_1 \\
\|\boldsymbol{w}\|_2 &= \sum_{i=1}^{n} \frac{\pm 1}{\sqrt{n}}^2 \\
&= \sum_{i=1}^{n} \frac{1}{n} \\
&= 1
\end{aligned}$$

By Cauchy-Schwarz, we get;

$$\sqrt{n}\|\boldsymbol{v}\|_1 = |\boldsymbol{v} \cdot \boldsymbol{w}| \leq \|\boldsymbol{v}\|_2\|\boldsymbol{w}\|_2 = \|\boldsymbol{v}\|_2$$

16th January 2020

Note that this recording has **no audio**, and therefore will just be the board transcribed. I honestly have no idea what he was doing in this lecture, it seems to just jump from topic to topic.

Equivalence of Norms?

Take any two norms on $\mathbb{R}^n$; $\| \ \|_a$, and $\| \ \|_b$.

$$\exists r, s \in \mathbb{R}^+ \ \forall \boldsymbol{v} \in \mathbb{R}^n \ [r\|\boldsymbol{v}\|_b \leq \|\boldsymbol{v}\|_a \leq s\|\boldsymbol{v}\|_b]$$

This means that norms in finite dimensional vector spaces are equivalent (no idea why, look it up).

Convergence of Vector Sequences

$(\boldsymbol{r}_n)$ is a sequence of vectors, and $(a_{i,j})$ is the i,j^{th} entry of $\boldsymbol{A}$. For a vector $\boldsymbol{v}^{(m)} \in \mathbb{R}^n$, where $m = 0, 1, 2, \ldots$

$$\boldsymbol{v}^{(m)} = \begin{bmatrix} v_1^{(m)} \\ v_2^{(m)} \\ \vdots \\ v_n^{(m)} \end{bmatrix}$$

For a vector sequence $\boldsymbol{v}^{(m)}$ to converge to some vector $\boldsymbol{v} \in \mathbb{R}^n$, the following must hold;

$$\boldsymbol{v}^{(m)} \to \boldsymbol{v} \in \mathbb{R}^n \Leftrightarrow \lim_{m\to\infty} \|\boldsymbol{v}^{(m)} - \boldsymbol{v}\| \to 0$$

This is componentwise convergence, such that $\forall i \in [1, n] \ [v_i^{(m)} \to v_i]$.

Matrix Norms

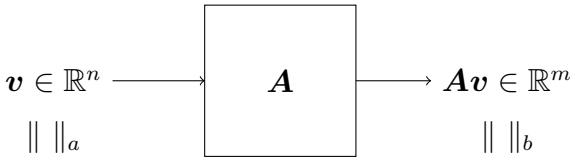
Vectors are a type of matrix. For a matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, the following properties of its norms must hold, where $\| \ \| : \mathbb{R}^{m\times n} \to \mathbb{R}_{\geq 0}$;

(i) $\|\boldsymbol{A}\| > 0$ given that $\boldsymbol{A} \neq \boldsymbol{0}$

(ii) $\|\lambda\boldsymbol{A}\| = |\lambda|\|\boldsymbol{A}\|$

(iii) $\|\boldsymbol{A} + \boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$

(iv) $\|\boldsymbol{B}\boldsymbol{A}\| \leq \|\boldsymbol{B}\|\|\boldsymbol{A}\|$



$$\|\boldsymbol{A}\boldsymbol{v}\|_b \leq \|\boldsymbol{A}\|\|\boldsymbol{v}\|_a$$

For the following example, take $(a_{i,j}) = \boldsymbol{A} \in \mathbb{R}^{m\times n}$;

$$a_j = \begin{bmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{m,j} \end{bmatrix} \qquad \text{the } j^{\text{th}} \text{ column of } \boldsymbol{A}$$

$$a^i = \begin{bmatrix} a_{i,1} & a_{i,2} & \cdots & a_{i,n} \end{bmatrix} \qquad \text{the } i^{\text{th}} \text{ row of } \boldsymbol{A}$$

We have the following norms on matrices;

$$\begin{aligned}
\|\boldsymbol{A}\|_1 &= \max\{\|a_j\|_1 : 1 \le j \le n\} \\
\|\boldsymbol{A}\|_\infty &= \max\{\|(a^i)^\top\|_1 : 1 \le i \le m\} \\
\|\boldsymbol{A}\|_F &= \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{i,j}|^2} && \text{Frobenius norm} \\
\|\boldsymbol{A}\|_2 &= \text{largest singular value of } \boldsymbol{A} \\
\text{let } \boldsymbol{A} &= \begin{bmatrix} 2 & 3 & 1 & 4 \\ 1 & 3 & -1 & 5 \\ \sqrt{2} & 0 & -2 & 2 \end{bmatrix} \\
\|\boldsymbol{A}\|_1 &= \max\{3+\sqrt{2}, 6, 4, 11\} \\
&= 11 \\
\|\boldsymbol{A}\|_\infty &= \max\{10, 10, 4+\sqrt{2}\} \\
&= 10 \\
\|\boldsymbol{A}\|_F &= \sqrt{4+9+1+16+1+9+1+25+2+0+4+4} \\
&= 2\sqrt{19}
\end{aligned}$$

Let there be two vector norms, $\|\ \|_a$ on $\mathbb{R}^n$ and $\|\ \|_b$ on $\mathbb{R}$. If $\|\ \|$ (matrix norm) satisfies

$$\forall \boldsymbol{A} \in \mathbb{R}^{m\times n}, x \in \mathbb{R}^n\ [\|\boldsymbol{A}\boldsymbol{x}\|_b \le \|\boldsymbol{A}\|\|\boldsymbol{x}\|_a]$$

then $\|\ \|$ is **consistent** with $\|\ \|_a$ and $\|\ \|_b$. Additionally if $a = b$, then $\|\ \|$ is **compatible** with $\|\ \|_a$. This gives us the following propositions;

- $\|\ \|_1$ (matrix norm) is compatible with $\|\ \|_1$ (vector norm)
- $\|\ \|_2$ (matrix norm) is compatible with $\|\ \|_2$ (vector norm)
- $\|\ \|_\infty$ (matrix norm) is compatible with $\|\ \|_\infty$ (vector norm)
- $\|\ \|_F$ (matrix norm) is compatible with $\|\ \|_2$ (vector norm) $\Rightarrow \|\boldsymbol{A}\boldsymbol{x}\|_2 \le \|\boldsymbol{A}\|_F\|\boldsymbol{x}\|_2$

Given a vector norm $\|\ \|$ on $\mathbb{R}^n$ then the matrix norm $\|\ \|$ subordinate to vector norm $\|\ \|$ is defined by

$$\|\boldsymbol{A}\| = \max\{\|\boldsymbol{A}\boldsymbol{x}\| : \|\boldsymbol{x}\| \le 1\} = \max\{\|\boldsymbol{A}\boldsymbol{x}\| : \|\boldsymbol{x}\| = 1\} = \max\{\|\boldsymbol{A}\tfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\| : \boldsymbol{x} \ne \boldsymbol{0}\}$$

Using this, we can prove property (iii) (see above). We claim that $\|\boldsymbol{A}+\boldsymbol{B}\| \le \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$ for matrix norm $\|\ \|$ subordinate to vector norm $\|\ \|$.

$$\begin{aligned}
\|\boldsymbol{A}+\boldsymbol{B}\| &= \max\{\|(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{x}\| : \|\boldsymbol{x}\| \le 1\} \\
&= \max\{\|\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{x}\| : \|\boldsymbol{x}\| \le 1\} \\
&\le \max\{\|\boldsymbol{A}\boldsymbol{x}\| + \|\boldsymbol{B}\boldsymbol{x}\| : \|\boldsymbol{x}\| \le 1\} && \text{triangle inequality for } \|\ \| \\
&\le \max\{\|\boldsymbol{A}\boldsymbol{x}\| : \|\boldsymbol{x}\| \le 1\} + \max\{\|\boldsymbol{B}\boldsymbol{x}\| : \|\boldsymbol{x}\| \le 1\} && \text{maximise independently} \\
&= \|\boldsymbol{A}\| + \|\boldsymbol{B}\| && \blacksquare
\end{aligned}$$

We are also able to prove property (iv), which we claim to be $\|\boldsymbol{B}\boldsymbol{A}\| \le \|\boldsymbol{B}\|\|\boldsymbol{A}\|$.

$$\begin{aligned}
\|\boldsymbol{B}\boldsymbol{A}\| &= \max\{\|\boldsymbol{B}\boldsymbol{A}\boldsymbol{x}\| : \|\boldsymbol{x}\| \le 1\} \\
\|\boldsymbol{B}\boldsymbol{A}\boldsymbol{x}\| &= \|\boldsymbol{B}(\boldsymbol{A}\boldsymbol{x})\| \\
\|\boldsymbol{A}\boldsymbol{x}\| &\le \|\boldsymbol{A}\|\|\boldsymbol{x}\| && \text{matrix norm subordinate to vector norm}
\end{aligned}$$

To show the line above, we consider the two cases, $\boldsymbol{x} \neq \boldsymbol{0}$ and $\boldsymbol{x} = \boldsymbol{0}$. If $\boldsymbol{x} = \boldsymbol{0}$, no work needs to be done, as it is trivial. I have no idea why this works, but he wrote it.

$$\begin{aligned}
\text{show } \left|\left|\boldsymbol{A}\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\right|\right| &\leq \|\boldsymbol{A}\| \\
\text{but } \left|\left|\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\right|\right| &= 1 \qquad \Rightarrow \\
\|\boldsymbol{A}\boldsymbol{x}\| &\leq \|\boldsymbol{A}\|\|\boldsymbol{x}\|
\end{aligned}$$

Continuing on, we have

$$\begin{aligned}
\|\boldsymbol{B}\boldsymbol{A}\boldsymbol{x}\| &= \|\boldsymbol{B}(\boldsymbol{A}\boldsymbol{x})\| \\
&\leq \|\boldsymbol{B}\|\|\boldsymbol{A}\boldsymbol{x}\| \\
&\leq \|\boldsymbol{B}\|\|\boldsymbol{A}\|\|\boldsymbol{x}\| \\
\|\boldsymbol{B}\boldsymbol{A}\| &= \max\{\|\boldsymbol{B}\boldsymbol{A}\boldsymbol{x}\| : \|\boldsymbol{x}\| \leq 1\} \\
&\leq \max\{\|\boldsymbol{B}\|\|\boldsymbol{A}\|\|\boldsymbol{x}\| : \|\boldsymbol{x}\| \leq 1\} \\
&= \|\boldsymbol{B}\|\|\boldsymbol{A}\| \max\{\|\boldsymbol{x}\| : \|\boldsymbol{x}\| \leq 1\} \qquad \text{obviously 1, as bounded on top} \\
&= \|\boldsymbol{B}\|\|\boldsymbol{A}\| \qquad \blacksquare
\end{aligned}$$

Complex Vectors

$$\mathbb{C}^n = \left\{ \boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} : v_i \in \mathbb{C} \right\}$$

$$\begin{aligned}
z \in \mathbb{C} &= a + ib \qquad a \in \mathbb{R}, b \in \mathbb{R}, i = \sqrt{-1} \\
z^* &= a - ib \\
|z| &= \sqrt{a^2 + b^2}
\end{aligned}$$

Take a linear map $f : \mathbb{C}^n \to \mathbb{C}^m$, the same properties hold;

$$f(a\boldsymbol{v} + b\boldsymbol{w}) = af(\boldsymbol{v}) + bf(\boldsymbol{w}) \qquad a, b \in \mathbb{C}$$

We also want to define something similar to the dot product in $\mathbb{R}^n$;

$$\begin{aligned}
\boldsymbol{v} \cdot \boldsymbol{w} &= \|\boldsymbol{v}\|_2\|\boldsymbol{w}\|_2 \cos\theta_{\boldsymbol{v},\boldsymbol{w}} \\
\boldsymbol{v} \cdot \boldsymbol{v} &= \sqrt{\|\boldsymbol{v}\|_2^2} \\
&= \|\boldsymbol{v}\|_2 \\
< \boldsymbol{v}, \boldsymbol{w} > &= \sum_{i=1}^{n} v_i^* w_i \\
< \boldsymbol{v}, \boldsymbol{v} > &= \sum_{i=1}^{n} v_i^* v_i \\
&= \sum_{i=1}^{n} |v_i|^2
\end{aligned}$$

The standard basis in $\mathbb{R}^n$ is defined as $(\boldsymbol{e_1}, \ldots, \boldsymbol{e_n})$ where

$$\boldsymbol{e_j} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} 1^{\text{st}} \\ 2^{\text{nd}} \\ \vdots \\ j^{\text{th}} \\ \vdots \\ n^{\text{th}} \end{matrix}$$

For any vector $\boldsymbol{v} \in \mathbb{C}^n$, it can be written in the standard basis as such;

$$\begin{aligned}\boldsymbol{v} &= \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \\ &= v_1\boldsymbol{e_1} + \cdots + v_n\boldsymbol{e_n}\end{aligned}$$

Basis Change, Again

Let the linear map $f : \mathbb{R}^n \to \mathbb{R}^m$.

$$\begin{aligned}\boldsymbol{B} &= (\boldsymbol{b_1}, \ldots, \boldsymbol{b_n}) && \text{an ordered basis of } \mathbb{R}^n \\ \boldsymbol{D} &= (\boldsymbol{d_1}, \ldots, \boldsymbol{d_m}) && \text{an ordered basis of } \mathbb{R}^m\end{aligned}$$

Find the matrix $\boldsymbol{A}$ ($\boldsymbol{A} := f_{\boldsymbol{DB}}$)representing f with respect to (?) $\boldsymbol{B}$ and $\boldsymbol{D}$.

$$\begin{aligned}\boldsymbol{v} &= \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} && \text{coordinates of a point } \boldsymbol{p} \in \mathbb{R}^n \\ \boldsymbol{p} &= \sum_{j=1}^{n} v_j \boldsymbol{b_j} \\ \boldsymbol{Av} & && \text{should be coordinate of } f(\boldsymbol{p}) \in \mathbb{R}^m \\ f(\boldsymbol{p}) &= \sum_{i=1}^{m} (\boldsymbol{Av})_i \boldsymbol{d_i} \\ f(\boldsymbol{b_j}) \in \mathbb{R}^m &= \sum_{i=1}^{m} a_{i,j}\boldsymbol{d_i} && 1 \leq j \leq n\end{aligned}$$

Take $\boldsymbol{B} \in \mathbb{R}^{m \times n}$, where $\boldsymbol{B} = [\boldsymbol{b_1}, \boldsymbol{b_2}, \ldots, \boldsymbol{b_n}]$, and $e_1, \ldots, \boldsymbol{e_m}$ being the standard basis (see previous).

$$\begin{aligned}\boldsymbol{Be_j} &= \boldsymbol{B}\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow j^{\text{th}} \\ &= b_{1,j}\boldsymbol{e_1} + ... + b_{m,j}\boldsymbol{e_m}\end{aligned}$$

Suppose $m = n$ and also $f = \text{id}$ (identity), but $\boldsymbol{B}$ and $\boldsymbol{D}$ are different. The matrix $(\text{id})_{\boldsymbol{DB}}$ represents a change of basis from $\boldsymbol{B}$ to $\boldsymbol{D}$. The point $\boldsymbol{p} \in \mathbb{R}^n$ has coordinates $\boldsymbol{x} \in \mathbb{R}^n$ with respect to $\boldsymbol{B}$, and $\boldsymbol{y} \in \mathbb{R}^n$ with respect to $\boldsymbol{D}$.

$$\boldsymbol{Dy} = \boldsymbol{p} = \boldsymbol{Bx} = x_1\boldsymbol{b_1} + \cdots + x_n\boldsymbol{b_n}$$

From this we gather $\boldsymbol{Dy} = \boldsymbol{Bx}$, therefore $\boldsymbol{y} = \boldsymbol{D}^{-1}\boldsymbol{Bx} = (\text{id})_{\boldsymbol{DB}}\boldsymbol{x}$, which means that

$$(\text{id})_{\boldsymbol{DB}} = \boldsymbol{D}^{-1}\boldsymbol{B}$$

Some stuff on functions between bases?

$$\begin{array}{ccccc} \mathbb{R}^n & \xrightarrow{\quad f \quad} & \mathbb{R}^m & \xrightarrow{\quad g \quad} & \mathbb{R}^k \\ \boldsymbol{B} & & \boldsymbol{C} & & \boldsymbol{D} \end{array}$$

$$\mathbb{R}^n_{\boldsymbol{B}} \xrightarrow{f_{\boldsymbol{CB}}} \mathbb{R}^m_{\boldsymbol{C}} \xrightarrow{g_{\boldsymbol{DC}}} \mathbb{R}^k_{\boldsymbol{D}}$$

$$g_{\boldsymbol{DC}} f_{\boldsymbol{CB}} = (g \circ f)_{\boldsymbol{DB}}$$

This then goes into change of basis, but see last year's **CO145**.

22nd January 2020

Tutorial Question

$f : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear map, and $\boldsymbol{E} = (\boldsymbol{e_1}, \boldsymbol{e_2})$ is an ordered basis.

$$f(\boldsymbol{e_1}) = 5\boldsymbol{e_1} - 6\boldsymbol{e_2}$$
$$f(\boldsymbol{e_2}) = 3\boldsymbol{e_1} + \boldsymbol{e_2}$$

We only care about what the linear map does to the ordered basis, as anything else can be done by linearity.

(i) Find $f_{\boldsymbol{EE}}$, the matrix representation of f in $\boldsymbol{E}$ - note that this has the same input space as the output space, but it can be different.

The first column can be done by reading the entry for for $f(\boldsymbol{e_1})$, and similarly for the second column as follows;

$$f_{\boldsymbol{EE}} = \begin{bmatrix} 5 & 3 \\ -6 & 1 \end{bmatrix}$$

(ii) If we have another ordered basis $\boldsymbol{D} = (\boldsymbol{d_1}, \boldsymbol{d_2})$, where $\boldsymbol{d_1} = \boldsymbol{e_1} - \boldsymbol{e_2}$ and $\boldsymbol{d_2} = \boldsymbol{e_1} + \boldsymbol{e_2}$, find $f_{\boldsymbol{DD}}$.

$$\begin{array}{ccc} \mathbb{R}^2_{\boldsymbol{E}} & \xrightarrow{f_{\boldsymbol{EE}}} & \mathbb{R}^2_{\boldsymbol{E}} \\ \Big\uparrow \boldsymbol{I_{ED}} & & \Big\downarrow \boldsymbol{I_{DE}} = (\boldsymbol{I_{ED}})^{-1} \\ \mathbb{R}^2_{\boldsymbol{D}} & \xrightarrow{f_{\boldsymbol{DD}}} & \mathbb{R}^2_{\boldsymbol{D}} \end{array}$$

$\boldsymbol{I_{ED}}$ can easily be obtained by reading the entries for $\boldsymbol{d_1}$ for the first column, and similarly for $\boldsymbol{d_2}$ in the second column;

$$\begin{aligned} \boldsymbol{I_{ED}} &= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ \boldsymbol{I_{DE}} &= (\boldsymbol{I_{ED}})^{-1} \\ &= \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \\ f_{\boldsymbol{DD}} &= \boldsymbol{I_{DE}} f_{\boldsymbol{EE}} \boldsymbol{I_{ED}} \\ &= \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 5 & 3 \\ -6 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 1 & 13 \\ -5 & 3 \end{bmatrix} \end{aligned}$$

Eigenvalues + Generalised Eigenvectors

Working with a matrix $\boldsymbol{A} \in \mathbb{C}^{m \times m}$. For an eigenvector $\boldsymbol{v} \in \mathbb{C}^m \setminus \{\boldsymbol{0}\}$, and an eigenvalue $\lambda \in \mathbb{C}$, $\boldsymbol{Av} = \lambda\boldsymbol{v} \Rightarrow (\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{v} = \boldsymbol{0} \Rightarrow |\boldsymbol{A} - \lambda\boldsymbol{I}| = 0 \Rightarrow P_{\boldsymbol{A}}(\lambda) = 0$ (characteristic polynomial). This complex polynomial will be of degree m, and it will have precisely m roots (including multiplicity). Suppose $P_{\boldsymbol{A}}(\lambda) = 0$, then $(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{v} = \boldsymbol{0}$ has a solution where $\boldsymbol{v} \neq \boldsymbol{0}$.

Assume we have $\lambda_1, \ldots, \lambda_t$ distinct eigenvalues, meaning that $P_{\boldsymbol{A}}(\lambda_i) = 0$ for $1 \leq i \leq t$. This means we can write the characteristic polynomial as;

$$P_{\boldsymbol{A}}(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \ldots (\lambda - \lambda_t)^{m_t}$$

Where m_i is the **algebraic multiplicity** of λ_i. $m_i \in \mathbb{N}$, and also $1 \leq m_i \leq m$, as it must not exceed the dimension of the matrix. On the other hand, the **geometric multiplicity** of λ_i is ℓ_i, which is the **nullity** of $(\boldsymbol{A} - \lambda_i \boldsymbol{I})$. The nullity is the dimension of the kernel / null-space. $1 \leq \ell_i \leq m_i$, as we already have at least one non-zero solution from $(\boldsymbol{A} - \lambda_i \boldsymbol{I})\boldsymbol{v} = \boldsymbol{0}$.

In the nice case, we have $\ell_i = m_i$, for $1 \leq i \leq t$, which means the matrix is diagonalisable. We have m_i linearly independent vectors $(\boldsymbol{v_{i,1}}, \boldsymbol{v_{i,2}}, \boldsymbol{\ldots}, \boldsymbol{v_{i,m_i}})$ which satisfy

$$\boldsymbol{A v_{i,j}} = \lambda_i \boldsymbol{v_{i,j}} \text{ for } 1 \leq j \leq m_i$$

If we take these eigenvectors as an ordered basis;

$$\boldsymbol{B} = [\boldsymbol{v_{1,1}}, \boldsymbol{v_{1,2}}, \ldots, \boldsymbol{v_{1,m_1}}, \boldsymbol{v_{2,1}}, \boldsymbol{v_{2,2}}, \ldots, \boldsymbol{v_{2,m_2}}, \ldots \boldsymbol{v_{t,1}}, \boldsymbol{v_{t,2}}, \ldots, \boldsymbol{v_{t,m_t}}]$$

We also want to note that $\sum_{i=1}^{t} m_i = m$, as that is the degree of the characteristic polynomial. Multiplying the basis by the original matrix, we get;

$$\boldsymbol{AB} = [\lambda_1 \boldsymbol{v_{1,1}}, \lambda_1 \boldsymbol{v_{1,2}}, \ldots, \lambda_1 \boldsymbol{v_{1,m_1}}, \lambda_2 \boldsymbol{v_{2,1}}, \lambda_2 \boldsymbol{v_{2,2}}, \ldots, \lambda_2 \boldsymbol{v_{2,m_2}}, \ldots \lambda_t \boldsymbol{v_{t,1}}, \lambda_t \boldsymbol{v_{t,2}}, \ldots, \lambda_t \boldsymbol{v_{t,m_t}}]$$

Since all the columns of $\boldsymbol{B}$ are linearly independently, by our definition, the inverse $\boldsymbol{B}^{-1}$ exists. Therefore, we can write

$$\boldsymbol{B}^{-1}\boldsymbol{AB} = \begin{bmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \lambda_1 & & & & & & \\ & & & \lambda_2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_2 & & & \\ & & & & & & \ddots & & \\ & & & & & & & \lambda_t & \\ & & & & & & & & \ddots \\ & & & & & & & & & \lambda_t \end{bmatrix} \text{ (everything else is 0)}$$

Which has m_1 instances of λ_1, followed by m_2 instances of λ_2, and so on, until m_t instances of λ_t.

Example for $\ell_i = m_i$

$$\boldsymbol{A} = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix}$$

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = (\lambda - 3)^2(\lambda - 5)$$

$\lambda_1 = 3$ — two linearly independently eigenvectors

$$\boldsymbol{v_{1,1}} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\boldsymbol{v_{1,2}} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$\lambda_2 = 5$

$$\boldsymbol{v_{2,1}} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$\boldsymbol{B} = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix}$$

$$\boldsymbol{B}^{-1} = \frac{1}{2}\begin{bmatrix} -2 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

Trivial Example for $\ell_i < m_i$

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$|\boldsymbol{A} - \lambda\boldsymbol{I}| = \lambda^2$$

$$\lambda_1 = 0$$

$$\boldsymbol{v_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \text{only solution, hence } \ell_1 = 1 < 2 = m_1$$

$$(\boldsymbol{A} - 0\boldsymbol{I}) \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

although this vector is not mapped to zero, it is mapped to something that **will** be mapped to zero

$$(\boldsymbol{A} - 0\boldsymbol{I})^2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = (\boldsymbol{A} - 0\boldsymbol{I})(\boldsymbol{A} - 0\boldsymbol{I}) \begin{bmatrix} 0 \\ 1 \end{bmatrix} = (\boldsymbol{A} - 0\boldsymbol{I}) \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \boldsymbol{0}$$

We say $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is a generalised eigenvector for $\lambda_1 = 0$. A vector which is not mapped by $(\boldsymbol{A} - \lambda\boldsymbol{I})$ to $\boldsymbol{0}$, but is $\boldsymbol{0}$ when iterated once more.

Less Trivial Example

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$|\boldsymbol{A} - \lambda\boldsymbol{I}| = (1 - \lambda)^3$$

$$\lambda_1 = 1 \qquad m_1 = 3$$

$$(\boldsymbol{A} - \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0} \qquad \Leftrightarrow$$

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \boldsymbol{x} = \boldsymbol{0}$$

this has rank 1, and therefore by rank-nullity theorem (rank + nullity = 3), has 2 linearly independent solutions, therefore $\ell_1 = 2 < 3 = m_1$

$$\boldsymbol{v_1} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

$$\boldsymbol{v_2} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

to find the generalised eigenvector $\boldsymbol{v_3}$, we want to find some vector that is mapped by $(\boldsymbol{A} - 1\boldsymbol{I})$ to the eigenspace, which is some linear combination of $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$

$$\begin{aligned}
(\boldsymbol{A} - \boldsymbol{I})\boldsymbol{v_3} &= \alpha_1\boldsymbol{v_1} + \alpha_2\boldsymbol{v_2} && \Leftrightarrow \\
\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} 0 \\ \alpha_1 \\ -\alpha_1 \end{bmatrix} + \begin{bmatrix} \alpha_2 \\ 0 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} x_2 + x_3 \\ 0 \\ 0 \end{bmatrix} && \Leftrightarrow \\
\alpha_1 &= 0 \\
x_2 + x_3 &= \alpha_2 && \text{let } x_2 = 0 \Rightarrow x_3 = \alpha_2 = 1 \\
\boldsymbol{v_3} &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\
\boldsymbol{B} &= [\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}] \\
&= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \\
\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} && \text{the extra 1 is from } \boldsymbol{v_2}?
\end{aligned}$$

This is in Jordan Normal Form. For some λ_i eigenvalue, and $\ell_i \leq m_i$, the sum of the sizes of the blocks is m_i, and the number of blocks is ℓ_i. If λ_i is an eigenvalue of $\boldsymbol{A}$ with algebraic multiplicity m_i, then the nullity of $(\boldsymbol{A} - \lambda_i\boldsymbol{I})^{m_i} = m_i$.

Definition: $\boldsymbol{v} \in \mathbb{R}^m$ is a generalised eigenvector for λ_i if $(\boldsymbol{A} - \lambda_i\boldsymbol{I})^{m_i}\boldsymbol{v} = \boldsymbol{0}$. The maximum iterations is m_i, but can be less.

23rd January 2020

In general;

- $\forall j\ [\ell_j = m_j] \Rightarrow \boldsymbol{A}$ is diagonalisable
- $\exists j\ [\ell_j < m_j] \Rightarrow \boldsymbol{A}$ is diagonalisable or **almost** diagonalisable

The following is a very poor explanation of what the form looks like, because I don't know how to draw blocks in LaTeX. See Panopto at timestamp 13:17 for a proper drawing.

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{B_{\lambda_1}} & & \\ & \ddots & \\ & & \boldsymbol{B_{\lambda_t}} \end{bmatrix}$$

$$\boldsymbol{B_j} = \begin{bmatrix} \boldsymbol{J_j} & & \\ & \ddots & \\ & & \boldsymbol{J_j} \end{bmatrix} \qquad \text{all } \boldsymbol{J_j} \text{ can be of different sizes, } \boldsymbol{B_j} \in \mathbb{R}^{m_j \times m_j}$$

$$\boldsymbol{J_i} = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

Spectral Decomposition of Symmetric Matrices

In almost all areas of computer science, symmetric matrices are the dominant paradigm. All the eigenvalues of real symmetric matrices are real, and are always diagonalisable. A symmetric matrix is always a square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, which also satisfies $\boldsymbol{A}^\top = \boldsymbol{A}$, meaning it is symmetric along the diagonal (the elements above the diagonal determine the elements below, and vice versa). We are to prove the following properties;

(i) All eigenvalues of $\boldsymbol{A}$ are real.

First assume that $\boldsymbol{v} \in \mathbb{C}^n$, as we don't yet know that it is real, therefore we can have the case $\lambda \in \mathbb{C}$.

$$\begin{aligned} \boldsymbol{A}\boldsymbol{v} &= \lambda \boldsymbol{v} && \Rightarrow \\ \boldsymbol{v}^{*\top} \boldsymbol{A}\boldsymbol{v} &= \lambda \boldsymbol{v}^{*\top} \boldsymbol{v} \\ &= \lambda \|\boldsymbol{v}\|^2 && (1) \\ \boldsymbol{v}^{*\top} \boldsymbol{A}\boldsymbol{v} &= \left(\boldsymbol{v}^{*\top} \boldsymbol{A}\right) \boldsymbol{v} \\ &= \left((\boldsymbol{A}\boldsymbol{v})^\top\right)^* \boldsymbol{v} \end{aligned}$$

briefly proving the above;

$$\begin{aligned} \left((\boldsymbol{A}\boldsymbol{v})^\top\right)^* &= \left(\boldsymbol{v}^\top \boldsymbol{A}^\top\right)^* && \text{taking the transpose inside the bracket} \\ &= \left(\boldsymbol{v}^\top\right)^* \boldsymbol{A}^* && \text{taking the conjugate inside and } \boldsymbol{A}^\top = \boldsymbol{A} \\ &= \left(\boldsymbol{v}^\top\right)^* \boldsymbol{A} && \boldsymbol{A} \text{ is real valued matrix} \end{aligned}$$

continuing on;

$$\begin{aligned} &= \left((\lambda \boldsymbol{v})^\top\right)^* \boldsymbol{v} && \boldsymbol{v} \text{ is an eigenvector} \\ &= \lambda^* \left(\boldsymbol{v}^\top\right)^* \boldsymbol{v} && \text{transposing a scalar does nothing} \\ &= \lambda^* \|\boldsymbol{v}\|^2 && (2) \end{aligned}$$

By comparing results, (1) = (2), hence $\lambda^* = \lambda$ (dividing through by $\|\boldsymbol{v}\|^2$ is allowed, since we know it is non-zero by properties of vector norms, and $\boldsymbol{v} \neq \boldsymbol{0}$ by properties of eigenvectors). As the complex conjugate of λ is the same as λ, we can conclude that $\lambda \in \mathbb{R}$.

(ii) Eigenvectors corresponding to different eigenvalues of $\boldsymbol{A}$ are perpendicular.

Note that with the above result, we know that $\boldsymbol{v} \in \mathbb{R}^n$, as there is no possible way to obtain complex values with a real valued matrix and a real valued eigenvalue.

$$\begin{aligned} \boldsymbol{A}\boldsymbol{v_1} &= \lambda_1 \boldsymbol{v_1} \\ \boldsymbol{A}\boldsymbol{v_2} &= \lambda_2 \boldsymbol{v_2} \\ \lambda_1 &\neq \lambda_2 \\ \boldsymbol{v_2}^\top \boldsymbol{A}\boldsymbol{v_1} &= \lambda_1 \boldsymbol{v_2}^\top \boldsymbol{v_1} \\ &= \lambda_1 (\boldsymbol{v_1} \cdot \boldsymbol{v_2}) && (1) \\ \boldsymbol{v_2}^\top \boldsymbol{A}\boldsymbol{v_1} &= \left(\boldsymbol{v_2}^\top \boldsymbol{A}\right) \boldsymbol{v_1} \\ &= (\boldsymbol{A}\boldsymbol{v_2})^\top \boldsymbol{v_1} && \text{adding a transpose and } \boldsymbol{A}^\top = \boldsymbol{A} \\ &= \lambda_2 \boldsymbol{v_2}^\top \boldsymbol{v_1} && \boldsymbol{v_2} \text{ is an eigenvector} \\ &= \lambda_2 (\boldsymbol{v_1} \cdot \boldsymbol{v_2}) && (2) \end{aligned}$$

Similarly, comparing results (1) and (2), we get

$$\begin{aligned} \lambda_1(\boldsymbol{v_1}\cdot\boldsymbol{v_2}) &= \lambda_2(\boldsymbol{v_1}\cdot\boldsymbol{v_2}) && \Leftrightarrow \\ \lambda_1(\boldsymbol{v_1}\cdot\boldsymbol{v_2}) - \lambda_2(\boldsymbol{v_1}\cdot\boldsymbol{v_2}) &= 0 \\ (\lambda_1-\lambda_2)(\boldsymbol{v_1}\cdot\boldsymbol{v_2}) &= 0 && \Leftrightarrow \\ \lambda_1 &= \lambda_2 \\ \text{or } \boldsymbol{v_1}\cdot\boldsymbol{v_2} &= 0 \end{aligned}$$

The former is not possible due to the condition that they are different eigenvalues, therefore $\boldsymbol{v_1}\cdot\boldsymbol{v_2}=0$. As neither are the zero vector, by definition of eigenvectors, $\cos\theta = 0$, therefore they are perpendicular.

(iii) (Tutorial 3) If λ_j is an eigenvalue of $\boldsymbol{A}$ then $\ell_j = m_j$, therefore $\boldsymbol{A}$ is diagonalisable.

Come back to this in two weeks?

To conclude, we get all the eigenvectors and we normalise them (divide by the ℓ_2-norm), and we can assume they are pairwise orthogonal. There is an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of $\mathbb{A}$; $(\boldsymbol{v_1},\ldots,\boldsymbol{v_n})$, where $\boldsymbol{v_j}\in\mathbb{R}^n$. This has the following properties;

- $\boldsymbol{A}\boldsymbol{v_j} = \lambda_j\boldsymbol{v_j}$
- $\boldsymbol{v_i}\perp\boldsymbol{v_j}$, when $i\neq j$ — perpendicular
- $\|\boldsymbol{v_i}\| = 1$ for $i = 1,\ldots,n$

Examples

For an example in $\mathbb{R}^{2\times2}$;

$$\begin{aligned} \boldsymbol{A} &= \begin{bmatrix}1 & 2\\ 2 & 1\end{bmatrix} \\ |\boldsymbol{A}-\lambda\boldsymbol{I}| &= (1-\lambda)^2 - 4 \\ \lambda_1 &= 3 \\ \lambda_2 &= -1 \\ \boldsymbol{v_1} &= \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} && \text{normalise straight away} \\ \boldsymbol{v_2} &= \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} \\ \boldsymbol{V} &= [\boldsymbol{v_1},\boldsymbol{v_2}] && \text{orthonormal basis of } \mathbb{R}^2 \\ &= \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} \\ \boldsymbol{AV} &= [\boldsymbol{A}\boldsymbol{v_1},\boldsymbol{A}\boldsymbol{v_2}] \\ &= [3\boldsymbol{v_1},-\boldsymbol{v_2}] \end{aligned}$$

For a symmetric matrix, if we have a matrix $\boldsymbol{V}$ representing the orthonormal basis of the eigenvectors, $\boldsymbol{V}^{-1} = \boldsymbol{V}^\top$

$$\begin{aligned} \boldsymbol{V}^{-1}\boldsymbol{AV} &= \boldsymbol{V}^\top\boldsymbol{AV} \\ &= \boldsymbol{V}^\top[3\boldsymbol{v_1},-\boldsymbol{v_2}] \\ &= \begin{bmatrix}\boldsymbol{v_1}^\top\\ \boldsymbol{v_2}^\top\end{bmatrix}[3\boldsymbol{v_1},-\boldsymbol{v_2}] \\ &= \begin{bmatrix}3 & 0\\ 0 & -1\end{bmatrix} && \text{dot product with unit length } (\boldsymbol{v_1}\perp\boldsymbol{v_2}) \end{aligned}$$

For an example in $\mathbb{R}^{3\times3}$;

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

now he just shows some intuition on getting an eigenvalue?

$$\begin{aligned}
|\boldsymbol{A} - \lambda\boldsymbol{I}| &= \begin{vmatrix} 1-\lambda & 1 & 3 \\ 1 & 3-\lambda & 1 \\ 3 & 1 & 1-\lambda \end{vmatrix} \\
&= \begin{vmatrix} 5-\lambda & 1 & 3 \\ 5-\lambda & 3-\lambda & 1 \\ 5-\lambda & 1 & 1-\lambda \end{vmatrix} && \boldsymbol{c_1}' = \boldsymbol{c_1} + \boldsymbol{c_2} + \boldsymbol{c_3} \\
&= 0 \\
\lambda_1 &= 5 \\
\text{trace}(\boldsymbol{A}) &= \lambda_1 + \lambda_2 + \lambda_3 \\
&= 5 && \Rightarrow \\
\lambda_2 + \lambda_3 &= 0 \\
|\boldsymbol{A}| &= \lambda_1\lambda_2\lambda_3 \\
&= -20 && \Rightarrow \\
\lambda_2\lambda_3 &= -4 \\
\lambda_2 &= 2 \\
\lambda_3 &= -2 \\
\boldsymbol{v_1} &= \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\
\boldsymbol{v_2} &= \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \\
\boldsymbol{v_3} &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \\
\boldsymbol{V} &= [\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}] \\
\boldsymbol{V}^{-1} &= \boldsymbol{V}^\top \\
\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V} &= \boldsymbol{V}^\top\boldsymbol{A}\boldsymbol{V} \\
&= \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix} && \text{spectral decomposition of a symmetric matrix}
\end{aligned}$$

Proof of $\boldsymbol{V}^{-1} = \boldsymbol{V}^\top$

We define $\boldsymbol{B} \in \mathbb{R}^{n\times n}$ as an **orthogonal matrix** if all columns $\boldsymbol{b_j}$ for $1 \leq j \leq n$ form an orthonormal basis of $\mathbb{R}^n$, such that the two properties hold;

- $\boldsymbol{b_j} \cdot \boldsymbol{b_i} = 0$ when $i \neq j$
- $\boldsymbol{b_j} \cdot \boldsymbol{b_j} = 1$

We want to prove that $\boldsymbol{B}^{-1} = \boldsymbol{B}^\top$:

$$
\begin{aligned}
\boldsymbol{B} &= [\boldsymbol{b_1}, \dots, \boldsymbol{b_n}] \\
\boldsymbol{B}^\top &= \begin{bmatrix} \boldsymbol{b_1}^\top \\ \vdots \\ \boldsymbol{b_n}^\top \end{bmatrix} \\
\boldsymbol{B}^\top \boldsymbol{B} &= \begin{bmatrix} \boldsymbol{b_1}^\top \\ \vdots \\ \boldsymbol{b_n}^\top \end{bmatrix} [\boldsymbol{b_1}, \dots, \boldsymbol{b_n}] \\
&= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \\
&= \boldsymbol{I}_n
\end{aligned}
$$

Preservation of Length

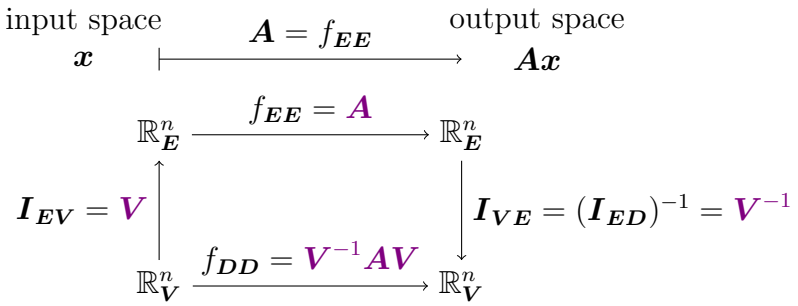
The length of a vector $\boldsymbol{x}$ is preserved under orthogonal transformation;

$$
\begin{aligned}
(\boldsymbol{V}\boldsymbol{x}) \cdot (\boldsymbol{V}\boldsymbol{x}) &= (\boldsymbol{V}\boldsymbol{x})^\top (\boldsymbol{V}\boldsymbol{x}) \\
&= \boldsymbol{x}^\top \boldsymbol{V}^\top \boldsymbol{V}\boldsymbol{x} \\
&= \boldsymbol{x}^\top \boldsymbol{V}^{-1} \boldsymbol{V}\boldsymbol{x} \\
&= \boldsymbol{x}^\top \boldsymbol{x} \\
&= \boldsymbol{x} \cdot \boldsymbol{x}
\end{aligned}
$$

Geometric Intuition

Take a real symmetric matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, such that $\boldsymbol{A}^\top = \boldsymbol{A}$. Assume we have processed it to get the eigenvalues $\lambda_1, \dots, \lambda_n$, such that $\boldsymbol{V}^\top \boldsymbol{A} \boldsymbol{V} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, and $\boldsymbol{V}$ is an orthonormal matrix formed of the eigenvectors.

Take a vector $\boldsymbol{x}$ from the input space, to $\boldsymbol{A}\boldsymbol{x}$ in the output space, mapped by $\boldsymbol{A}$;



Consider a unit sphere, created by $\|\boldsymbol{x}\| = 1$, where $\boldsymbol{x} \in \mathbb{R}^3$. The transformations $\boldsymbol{V}^{-1} = \boldsymbol{V}^\top$ and $\boldsymbol{V}$ are rotations as they preserve length. Take a point $\boldsymbol{x}$ on the surface of this unit sphere, applying $\boldsymbol{V}^{-1}$ to it simply rotates it to another point on the surface of the sphere (going from $\mathbb{R}^3_{\boldsymbol{E}}$ to $\mathbb{R}^3_{\boldsymbol{V}}$) in the input space. Let this point have coordinates $\boldsymbol{u}$ with respect to $\boldsymbol{V}$. Under the diagonal map $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$, we get the following;

$$u_1 \boldsymbol{v_1} + \cdots + u_n \boldsymbol{v_n} \mapsto \lambda_1 u_1 \boldsymbol{v_1} + \cdots + \lambda_n u_n \boldsymbol{v_n}$$

As $\boldsymbol{u}$ is on the sphere, we can say;

$$\|\boldsymbol{u}\|_2^2 = u_1^1 + \cdots + u_n^2 = 1$$

Therefore, the mapped point $\boldsymbol{y}$ can be written as $\begin{bmatrix} \lambda_1 u_1 \\ \vdots \\ \lambda_n u_n \end{bmatrix}$, satisfying (in three dimensions);

$$\frac{y_1^2}{\lambda_1^2} + \frac{y_2^2}{\lambda_2^2} + \frac{y_3^2}{\lambda_3^2} = 1$$

This gives a locus of an ellipsoid.

Rank-Nullity Theorem

Take a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$.

$$\begin{aligned} \operatorname{Im}(\boldsymbol{A}) \subseteq \mathbb{R}^m &= \{\boldsymbol{Ax} : \boldsymbol{x} \in \mathbb{R}^n\} \\ \dim(\operatorname{Im}(\boldsymbol{A})) &= \text{column rank of } \boldsymbol{A} \\ \operatorname{Null}(\boldsymbol{A}) \subseteq \mathbb{R}^n &= \{\boldsymbol{v} : \boldsymbol{Av} = \boldsymbol{0}\} && \text{same as kernel} \\ \text{rk(column space)} &= \text{rk(row space)} \\ \text{row space } \boldsymbol{A} &= \text{column space } \boldsymbol{A}^\top \end{aligned}$$

For elementary row operations, we have the following property;

(I) they are invertible.

If there is one of the three elementary row operations R that takes $\boldsymbol{A} \rightsquigarrow \boldsymbol{B}$, then $\exists R^{-1}$ that takes $\boldsymbol{B} \rightsquigarrow \boldsymbol{A}$.

If $\boldsymbol{A} \rightsquigarrow \boldsymbol{B}$, then $\boldsymbol{B} = \boldsymbol{M_R A}$. We obtain $\boldsymbol{M_R}$ by applying R to $\boldsymbol{I}_m$, such that $\boldsymbol{I}_m \rightsquigarrow \boldsymbol{M_R}$.

A reduction to row echelon form can be written as follows (with $-$ as a pivot, and $\times$ being any number);

$$\boldsymbol{A} \overset{R_1}{\rightsquigarrow} \cdots \overset{R_t}{\rightsquigarrow} \boldsymbol{C} = \left[\begin{array}{cccccc} - & \times & \times & \times & \times & \times \\ 0 & - & \times & \times & \times & \times \\ 0 & 0 & 0 & - & \times & \times \\ 0 & 0 & 0 & 0 & - & \times \\ 0 & 0 & 0 & 0 & 0 & - \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

In general it creates a staircase of the pivots, and everything below it is 0. Note that the example above is just a random example, as long as the general structure remains, it's fine.

An ERO doesn't change the row space, as it is operating on the rows. Swapping rows, multiplying by a non-zero number, and adding rows will not change anything. The claim is that while the column space can change, the rank of the column space doesn't change.

$$\begin{aligned} \boldsymbol{A} &= [\boldsymbol{a_1}, \ldots, \boldsymbol{a_n}] \\ \boldsymbol{A} \overset{R}{\rightsquigarrow} \boldsymbol{M_R A} &= [\boldsymbol{M_R a_1}, \ldots, \boldsymbol{M_R a_n}] \\ \boldsymbol{M_{R^{-1}}} &= (\boldsymbol{M_R})^{-1} && \text{therefore } \boldsymbol{M_R} \text{ is invertible} \end{aligned}$$

taking some linear combination of the vectors in $\boldsymbol{A}$

$$\begin{aligned} \lambda_1 \boldsymbol{a_{i_1}} + \cdots + \lambda_1 \boldsymbol{a_{i_t}} &= 0 && \Leftrightarrow \\ \boldsymbol{M_R}(\lambda_1 \boldsymbol{a_{i_1}} + \cdots + \lambda_1 \boldsymbol{a_{i_t}}) &= 0 \end{aligned}$$

This is due to the fact that multiplying a set of linearly independent vectors by an invertible matrix preserves independence. From this it follows that the dimension of the columns of $\boldsymbol{A}$ is equal to the dimension of the columns of $\boldsymbol{M_R A}$, since $\boldsymbol{M_R}$ is invertible. Therefore neither the row rank nor the column rank changes.

Since $\boldsymbol{C}$ is in REF, the row rank is equal to the column rank, which is equal to the number of pivots, let it be r, then $\text{rk}(\boldsymbol{A}) = r$.

29th January 2020

Properties of Symmetric Matrices

Take $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, such that $\boldsymbol{A}^\top = \boldsymbol{A}$.

All eigenvalues of $\boldsymbol{A}$ are positive (non-negative) iff $\forall \boldsymbol{x} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ $[\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} > 0 (\geq 0)]$. $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ is called a quadratic form. For example;

$$
\begin{aligned}
\boldsymbol{A} &= \begin{bmatrix} a & b \\ b & c \end{bmatrix} \\
\boldsymbol{x} &= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&= ax_1^2 + 2bx_1x_2 + cx_2^2 \\
&= k \\
\text{assume } x_2 &\neq 0 \\
\frac{\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}}{x_2^2} &= ay^2 + 2b^y + c && \text{where } y = \frac{x_1}{x_2} \\
&> 0 && \text{focusing on the positive case}
\end{aligned}
$$

If $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} > 0$, then we have an ellipse, with slanted major / minor axes. This uses a similar result to the result with the sphere's orthonormal basis being formed from the eigenvectors.

Considering this algebraically, if the quadratic in y is always greater than 0, then it must have no roots, hence the discriminant must be negative. Therefore we have the case where $(2b)^2 - 4ac < 0$, so $b^2 - ac < 0$, and $a > 0$.

Using the properties of the trace and the determinant, we can state the following;

$$
\begin{aligned}
\lambda_1 + \lambda_2 &= a + c \\
\lambda_1 \lambda_2 &= ac - b^2
\end{aligned}
$$

We want both $ac - b^2$ and $a + b$ to be positive. The former is shown with the result for the discriminant, as if $b^2 - ac$ is negative, then $ac - b^2$ must be positive. The latter then follows as c is positive only if a is positive (due to $ac > b^2 \geq 0$), which we have previously. This shows it holds for the simple case.

The proof for any $n > 1$ is as follows;

$$
\begin{aligned}
\boldsymbol{S} &= \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \\
&= \boldsymbol{V}^\top \boldsymbol{A}\boldsymbol{V} && \Rightarrow \\
\boldsymbol{A} &= \boldsymbol{V}\boldsymbol{S}\boldsymbol{V}^\top && \Rightarrow \\
\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} &= \boldsymbol{x}^\top \boldsymbol{V}\boldsymbol{S}\boldsymbol{V}^\top \boldsymbol{x} \\
&= \boldsymbol{z}^\top \boldsymbol{S}\boldsymbol{z} && \text{where } \boldsymbol{z} = \boldsymbol{V}^\top \boldsymbol{x} (\text{which preserves } \ell_2\text{-norm}) \\
&= \sum_{i=1}^{n} \lambda_i z_i^2
\end{aligned}
$$

To say $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ is positive for any non-zero $\boldsymbol{x}$ is to say that $\sum_{i=1}^{n} \lambda_i z_i^2$ is positive for any non-zero $\boldsymbol{v}$.

Then obviously all the eigenvalues (λ_i) must be positive iff $\sum_{i=1}^{n} \lambda_i z_i^2$ is positive.

Singular Value Decomposition (SVD)

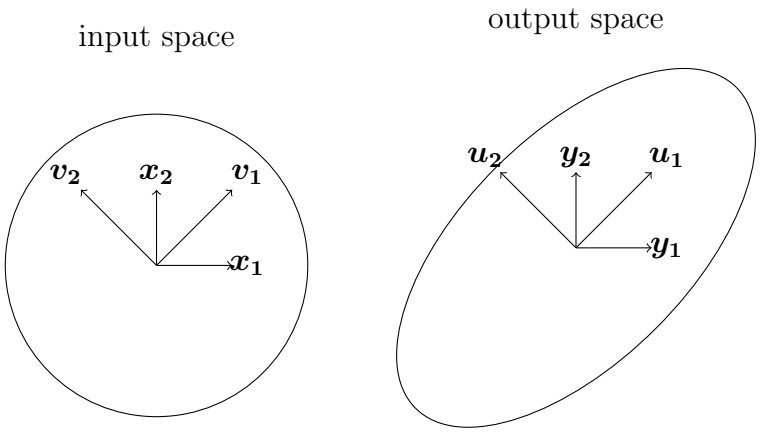
This works for a **general** matrix (not necessarily square). Take any matrix $\mathbb{R}^{m\times n}$.

$$\exists \boldsymbol{V} \in \mathbb{R}^{n\times n}, \boldsymbol{U} \in \mathbb{R}^{m\times m}, \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_p = 0$$

Where $\boldsymbol{V}$, $\boldsymbol{U}$ are orthogonal matrices, $p = \min(m, n)$, and $r = \text{rk}(\boldsymbol{A})$. We have a finitely decreasing sequence of σ. This allows us to write the following (assuming $n \leq m$);

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^\top, \text{ where } \boldsymbol{S} = \left[\begin{array}{cccc} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \\ \hline 0 & & \cdots & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & \cdots & & 0 \end{array}\right]$$

This is the singular value decomposition of $\boldsymbol{A}$, and $\sigma_1, \ldots, \sigma_r$ are the singular values of $\boldsymbol{A}$. σ_1 is the largest singular value of $\boldsymbol{A}$, which is the ℓ_2 matrix norm of $\boldsymbol{A}$ ($\|\boldsymbol{A}\|_2$). The geometric intuition of this, in two dimensions, is as follows;



Starting with our coordinates in the basis $(\boldsymbol{x_1}, \boldsymbol{x_2})$, the matrix $\boldsymbol{V}^\top$ performs a rotation to the new basis $(\boldsymbol{v_1}, \boldsymbol{v_2})$.

Take some point $\boldsymbol{x}$ on the unit circle with coordinates $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, we obtain the new coordinates $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \boldsymbol{V}^\top \boldsymbol{x}$.

We still maintain $\|\boldsymbol{x}\| = 1$ and $\|\boldsymbol{v}\| = 1$.

We're now in the basis of $(\boldsymbol{u_1}, \boldsymbol{u_2})$. We can say that $u_1 = \sigma_1 v_1$, $u_2 = \sigma_2 v_2$ (and so on, for higher dimensions). This implies the following result;

$$\frac{u_1^2}{\sigma_1^2} + \frac{u_2^2}{\sigma_2^2} = v_1^2 + v_2^2 = 1$$

Hence we obtain an ellipse.

Example

$$
\begin{aligned}
\boldsymbol{A} \in \mathbb{R}^{3\times 2} &= \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix} \\
\operatorname{rk}(\boldsymbol{A}) &= 1 \\
\boldsymbol{A} &= \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^\top \\
\boldsymbol{S} \in \mathbb{R}^{3\times 2} &= \operatorname{diag}(\sigma_1, \sigma_2, \dots) \\
\boldsymbol{A}^\top &= \left(\boldsymbol{V}^\top\right)^\top \boldsymbol{S}^\top \boldsymbol{U}^\top \\
&= \boldsymbol{V}\boldsymbol{S}^\top \boldsymbol{U}^\top \\
\boldsymbol{A}^\top\boldsymbol{A} &= \boldsymbol{V}\boldsymbol{S}^\top \boldsymbol{U}^\top \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^\top \\
&= \boldsymbol{V}\boldsymbol{S}^\top \boldsymbol{S}\boldsymbol{V}^\top && \boldsymbol{U} \text{ is an orthogonal matrix}
\end{aligned}
$$

brief note on $\boldsymbol{S}^\top\boldsymbol{S} \in \mathbb{R}^{2\times 2}$

$$
\begin{aligned}
\boldsymbol{S}^\top\boldsymbol{S} &= \operatorname{diag}(\sigma_1, \sigma_2, \dots)\ \operatorname{diag}(\sigma_1, \sigma_2, \dots) \\
&= \operatorname{diag}(\sigma_1^2, \sigma_2^2, \dots)
\end{aligned}
$$

brief proof for symmetry of $\boldsymbol{A}^\top\boldsymbol{A}$

$$
\begin{aligned}
\left(\boldsymbol{A}^\top\boldsymbol{A}\right)^\top &= \boldsymbol{A}^\top\left(\boldsymbol{A}^\top\right)^\top \\
&= \boldsymbol{A}^\top\boldsymbol{A}
\end{aligned}
$$

$\boldsymbol{A}^\top\boldsymbol{A}$ is also positive semi-definite, meaning all its eigenvalues are non-negative (≥ 0) since

$$
\begin{aligned}
\boldsymbol{x}^\top\boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{x} &= (\boldsymbol{A}\boldsymbol{x})^\top(\boldsymbol{A}\boldsymbol{x}) \\
&\geq 0
\end{aligned}
$$

In this form, all we have to do is to find the spectral decomposition for $\boldsymbol{A}^\top\boldsymbol{A}$, since we have a symmetric matrix.

$$
\begin{aligned}
\boldsymbol{A}^\top\boldsymbol{A} &= \begin{bmatrix} 3 & -3 \\ -3 & 3 \end{bmatrix} \\
|\boldsymbol{A}^\top\boldsymbol{A} - \lambda\boldsymbol{I}| &= (3-\lambda)^2 - 9 \\
\lambda_1 &= 6 && \Rightarrow \\
\sigma_1 &= \sqrt{6} \\
\boldsymbol{v_1} &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \\
\lambda_2 &= 0 && \Rightarrow \\
\sigma_2 &= 0 \\
\boldsymbol{v_2} &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
\boldsymbol{V} &= [\boldsymbol{v_1}, \boldsymbol{v_2}] \\
&= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\
\boldsymbol{A}\boldsymbol{V} &= \boldsymbol{U}\boldsymbol{S} && \Rightarrow \\
[\boldsymbol{A}\boldsymbol{v_1}, \boldsymbol{A}\boldsymbol{v_2}] &= \boldsymbol{U}\begin{bmatrix} \sqrt{6} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \\
&= [\sqrt{6}\boldsymbol{u_1}, \boldsymbol{0}] && \Rightarrow
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{u_1} &= \frac{1}{\sqrt{6}}\boldsymbol{A}\boldsymbol{v_1} \\
&= \frac{1}{\sqrt{3}}\begin{bmatrix}1\\-1\\1\end{bmatrix} \\
\boldsymbol{u_2} &= \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} && \text{unit vector orthogonal to } \boldsymbol{u_1} \\
\boldsymbol{u_3} &= \boldsymbol{u_1} \times \boldsymbol{u_2} \\
&= \frac{1}{\sqrt{6}}\begin{bmatrix}1\\-1\\-2\end{bmatrix} \\
\boldsymbol{A} &= [\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}]\begin{bmatrix}\sqrt{6} & 0\\ 0 & 0\\ 0 & 0\end{bmatrix}[\boldsymbol{v_1}, \boldsymbol{v_2}]
\end{aligned}
$$

30th January 2020

Note that this lecture has no audio recording for some of the first part. To recap; the goal of SVD is to find the following, such that $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^\top$

- $\boldsymbol{A} \in \mathbb{R}^{m\times n}$
- $\boldsymbol{U} \in \mathbb{R}^{m\times m}$ $\qquad \boldsymbol{U}^\top = \boldsymbol{U}$
- $\boldsymbol{V} \in \mathbb{R}^{n\times n}$ $\qquad \boldsymbol{V}^\top = \boldsymbol{V}$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_p = 0$ $\qquad p = \min\{m, n\},\ r = \mathrm{rk}(\boldsymbol{A})$
- $\boldsymbol{S} \in \mathbb{R}^{m\times n} = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r, 0, \ldots, 0)$

Example

While normally $n \leq m$, we have the following example for $m < n$, where we work with $\boldsymbol{A}\boldsymbol{A}^\top \in \mathbb{R}^{m\times m}$;

$$
\begin{aligned}
\boldsymbol{A} &= \begin{bmatrix}3 & 2 & 2\\ 2 & 3 & -2\end{bmatrix} \\
\boldsymbol{A}^\top &= \begin{bmatrix}3 & 2\\ 2 & 3\\ 2 & -2\end{bmatrix} \\
\boldsymbol{A}\boldsymbol{A}^\top &= \boldsymbol{U}\boldsymbol{S}\boldsymbol{S}^\top\boldsymbol{U}^\top \\
\boldsymbol{A}\boldsymbol{A}^\top &= \begin{bmatrix}17 & 8\\ 8 & 17\end{bmatrix} \\
\lambda_1 + \lambda_2 &= \text{trace} \\
&= 34 \\
\lambda_1\lambda_2 &= \text{determinant} \\
&= 225 \\
\lambda_1 &= 25 \\
\sigma_1 &= 5 \\
\lambda_2 &= 9 \\
\sigma_2 &= 3 \\
\boldsymbol{u_2} &= \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} \\
\boldsymbol{u_1} &= \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} && \text{orthogonal to } \boldsymbol{u_2}
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{U} &= \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} \\
\boldsymbol{A} &= \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^\top && \Rightarrow \\
\boldsymbol{A}^\top &= \boldsymbol{V}\boldsymbol{S}^\top\boldsymbol{U}^\top && \Rightarrow \\
\boldsymbol{A}^\top\boldsymbol{U} &= \boldsymbol{V}\boldsymbol{S}^\top && \Rightarrow \\
[\boldsymbol{A}^\top\boldsymbol{u_1}, \boldsymbol{A}^\top\boldsymbol{u_2}] &= [5\boldsymbol{v_1}, 3\boldsymbol{v_2}] && \Rightarrow \\
\boldsymbol{v_1} &= \frac{1}{5}\boldsymbol{A}^\top\boldsymbol{u_1} \\
&= \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\end{bmatrix} \\
\boldsymbol{v_2} &= \frac{1}{3}\boldsymbol{A}^\top\boldsymbol{u_2} \\
&= \frac{1}{3\sqrt{2}}\begin{bmatrix}1\\-1\\4\end{bmatrix} \\
\boldsymbol{v_3} &= \boldsymbol{v_1}\times\boldsymbol{v_2} && \text{orthogonal to both} \\
&= \frac{1}{3}\begin{bmatrix}2\\-2\\-1\end{bmatrix}
\end{aligned}
$$

General Strategy

$$
\underbrace{\boldsymbol{A}^\top\boldsymbol{A}}_{\substack{\text{symmetric}\\ \text{+ve semi-definite}}} = \boldsymbol{V}\boldsymbol{S}^\top\underbrace{\boldsymbol{U}^\top\boldsymbol{U}}_{\boldsymbol{I}}\boldsymbol{S}\boldsymbol{V}^\top = \boldsymbol{V}\underbrace{\boldsymbol{S}^\top\boldsymbol{S}}_{\in\mathbb{R}^{n\times n}}\boldsymbol{V}^\top
$$

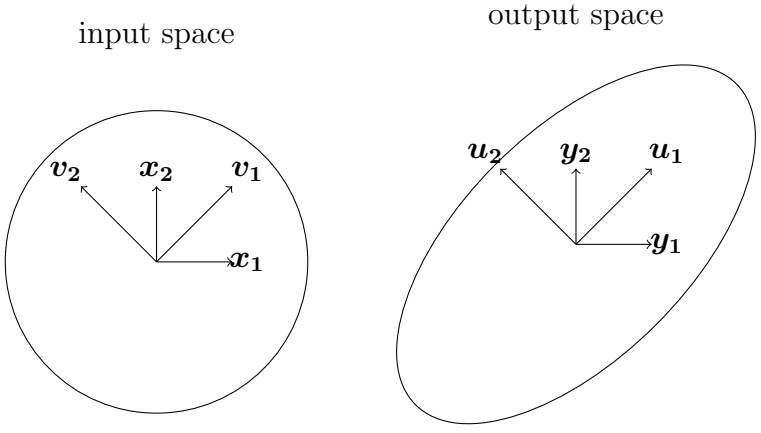
Solve for the spectral decomposition of $\boldsymbol{A}^\top\boldsymbol{A}$ to get $\boldsymbol{V}$, and $\lambda_1 = \sigma_1^2, \lambda_2 = \sigma_2^2, \ldots, \lambda_r = \sigma_r^2$. From here we can do the following;

$$
\begin{aligned}
\boldsymbol{A} &= \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^\top && \Rightarrow \\
\boldsymbol{A}\boldsymbol{V} &= \boldsymbol{U}\boldsymbol{S} && \Rightarrow \\
\boldsymbol{A}\boldsymbol{v_i} &= \sigma_i\boldsymbol{u_i} && (1 \le i \le r), \Rightarrow \\
\boldsymbol{u_i} &= \frac{\boldsymbol{A}\boldsymbol{v_i}}{\sigma_i}
\end{aligned}
$$

For the remaining vectors, $\boldsymbol{u_{r+1}}, \ldots, \boldsymbol{u_n}$ find vectors to form an orthonormal basis. Use systems of linear equations in dimensions higher than 3, if the dimension is 3, just use the cross product. Solve $\boldsymbol{u_{r+1}}$ to $\boldsymbol{u_n}$ with

$$
\boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{u_j} = \boldsymbol{0} \text{ for } r+1 \le j \le n
$$

such that $(\boldsymbol{u_1}, \ldots, \boldsymbol{u_r}, \boldsymbol{u_{r+1}}, \ldots, \boldsymbol{u_n})$ forms an orthonormal basis.



Using the same example as before, because I'm too lazy to draw this again, we see that $\boldsymbol{v_i} \mapsto \sigma_i \boldsymbol{v_i} = \boldsymbol{u_i}$. $\boldsymbol{v_1}$ is called the **principal component**, as it is the direction of the biggest change (since we have an ordering on σ).

A brief note on why $\boldsymbol{S}\boldsymbol{S}^\top$ gives a square matrix;

$$\boldsymbol{S} = \left[\begin{array}{ccc} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ \hline 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{array}\right]$$

$$\boldsymbol{S}^\top = \left[\begin{array}{ccc|ccc} \sigma_1 & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & \ddots & \vdots \\ & & \sigma_r & 0 & \cdots & 0 \end{array}\right]$$

$$\begin{aligned} \boldsymbol{S}^\top \boldsymbol{S} &= \boldsymbol{S}\boldsymbol{S}^\top && \text{by symmetry} \\ &= \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_r^2 \end{bmatrix} \end{aligned}$$

Cholesky Factorisation / Decomposition of Symmetric Positive Definite Matrices

Take a matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, with $\boldsymbol{A} = \boldsymbol{A}^\top$ and $\boldsymbol{A}$ is positive (semi-)definite. Then the following holds

$$\exists \boldsymbol{L} \in \mathbb{R}^{n \times n} \left[\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^\top\right]$$

where $\boldsymbol{L}$ is a lower triangular matrix, with all diagonal elements being positive (non-zero);

$$\boldsymbol{L} = \begin{bmatrix} l_{1,1} & 0 & \cdots & 0 \\ \times & l_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \times & \cdots & \times & l_{n,n} \end{bmatrix}$$

This is a general result to the following, in the real numbers; $a > 0 \Rightarrow \exists l > 0 \; [a = l^2]$. This algorithm is done with brute force, by setting $\boldsymbol{A} = (a_{i,j})$, $\boldsymbol{L} = (l_{i,j})$, and $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^\top$. The following is a simple example of this algorithm, with a 3×3 matrix;

$$\begin{aligned} \boldsymbol{A} &= \begin{bmatrix} 1 & -1 & 1 \\ -1 & 10 & -1 \\ 1 & -1 & 5 \end{bmatrix} \\ &= \boldsymbol{L}\boldsymbol{L}^\top \\ \boldsymbol{L} &= \begin{bmatrix} l_{1,1} & 0 & 0 \\ l_{2,1} & l_{2,2} & 0 \\ l_{3,1} & l_{3,2} & l_{3,3} \end{bmatrix} \\ \boldsymbol{L}^\top &= \begin{bmatrix} l_{1,1} & l_{2,1} & l_{3,1} \\ 0 & l_{2,2} & l_{3,2} \\ 0 & 0 & l_{3,3} \end{bmatrix} \end{aligned}$$

by multiplying out $\boldsymbol{L}\boldsymbol{L}^\top$, we can solve per column, starting with the first

$$\begin{aligned}
l_{1,1}^2 &= 1 && \Rightarrow \\
l_{1,1} &= 1 && \text{taking positive square root} \\
l_{2,1}l_{1,1} &= -1 && \Rightarrow \\
l_{2,1} &= -1 \\
l_{3,1}l_{1,1} &= 1 && \Rightarrow \\
l_{3,1} &= 1
\end{aligned}$$

second column

$$\begin{aligned}
l_{2,1}^2 + l_{2,2}^2 &= 10 && \Rightarrow \\
l_{2,2} &= 3 \\
l_{3,1}l_{2,1} + l_{3,2}l_{2,2} &= -1 && \Rightarrow \\
-1 + 3l_{3,2} &= -1 && \Rightarrow \\
l_{3,2} &= 0
\end{aligned}$$

third (final) column

$$\begin{aligned}
l_{3,1}^2 + l_{3,2}^2 + l_{3,3}^2 &= 5 && \Rightarrow \\
1 + 0 + l_{3,3}^2 &= 5 && \Rightarrow \\
l_{3,3} &= 2
\end{aligned}$$

setting the matrix

$$\boldsymbol{L} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

This can be used to easily solve the following;

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b} && \Leftrightarrow \\
\boldsymbol{L}\boldsymbol{L}^\top\boldsymbol{x} &= \boldsymbol{b} && \Leftrightarrow \\
\boldsymbol{L}\boldsymbol{y} &= \boldsymbol{b} && \text{let } \boldsymbol{y} = \boldsymbol{L}^\top\boldsymbol{x}
\end{aligned}$$

This is trivial to solve with forward substitution (y_1 can be solved easily, then substituted into the next equation to get y_2 and so on) as we have a lower triangular matrix

$$\begin{aligned}
b_1 &= l_{1,1}y_1 + 0 + \cdots + 0 \\
b_2 &= l_{2,1}y_1 + l_{2,2}y_2 + 0 + \cdots + 0 \\
b_3 &= l_{3,1}y_1 + l_{3,2}y_2 + l_{3,3}y_3 + 0 + \cdots + 0 \\
&\vdots \\
b_n &= l_{n,1}y_1 + \cdots + l_{n,n}y_n
\end{aligned}$$

If $\boldsymbol{L}$ is lower triangular, then $\boldsymbol{L}^\top$ must be upper triangular, and thus can be solved with backwards substitution (such that x_n is used to solve x_{n-1}, and so on)

$$\begin{aligned}
\boldsymbol{U} &= \boldsymbol{L}^\top \\
y_n &= 0 + \cdots + 0 + u_{n,n}x_n \\
y_{n-1} &= 0 + \cdots + 0 + u_{n-1,n-1}x_{n-1} + u_{n-1,n}x_n \\
&\vdots \\
y_1 &= u_{1,1}x_1 + \cdots + u_{1,n}x_n
\end{aligned}$$

Rank-Nullity Theorem, Again

For a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, the nullity of $\boldsymbol{A}$ (dimension of the kernel of $\boldsymbol{A}$) is $n - \text{rk}(\boldsymbol{A})$. To recap, a reduction to reduced row echelon form (RREF) is being in row echelon form, with the additional properties that all the pivots must be 1, and all other numbers in a column with a pivot are 0.

Applying the rank-nullity theorem to $\boldsymbol{A}^\top$, where $\boldsymbol{A}^\top \in \mathbb{R}^{n \times m}$;

$$\begin{aligned} \text{nullity}(\boldsymbol{A}^\top) + \text{rk}(\boldsymbol{A}^\top) = m & \qquad \Rightarrow \\ \dim(\ker(\boldsymbol{A}^\top)) + \text{rk}(\boldsymbol{A}) = m & \qquad \text{rk}(\boldsymbol{A}) = \text{rk}(\boldsymbol{A}^\top), \Rightarrow \\ \dim(\ker(\boldsymbol{A}^\top)) + \dim(\text{Im}(\boldsymbol{A})) = m & \qquad \text{rk}(\boldsymbol{A}) = \dim(\text{Im}(\boldsymbol{A})) \end{aligned}$$

It's important to now note that these are both subspaces of $\mathbb{R}^m$

$$\begin{aligned} \text{Im}(\boldsymbol{A}) \subseteq \mathbb{R}^m & \qquad \text{output space of } \boldsymbol{A} \\ \ker(\boldsymbol{A}^\top) \subseteq \mathbb{R}^m & \end{aligned}$$

Working through an example matrix, we can observe the following;

$$\begin{aligned} \boldsymbol{A} &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & -1 \end{bmatrix} \\ \boldsymbol{A}^\top &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \\ \text{Im}(\boldsymbol{A}) &= \left\{ x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} : x_1, x_2 \in \mathbb{R} \right\} \end{aligned}$$

geometrically, $\text{Im}(\boldsymbol{A})$ is a plane going through the origin - intuitively, we can represent this with a normal (take cross product)

$$\begin{aligned} &= \left\{ \boldsymbol{v} : \boldsymbol{v} \cdot \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix} = 0 \right\} \\ \ker(\boldsymbol{A}^\top) &= \{ \boldsymbol{w} : \boldsymbol{A}^\top \boldsymbol{w} = \boldsymbol{0} \} \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} &= \boldsymbol{0} \qquad \Rightarrow \\ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \boldsymbol{w} &= 0 \\ \text{and} \quad \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \cdot \boldsymbol{w} &= 0 \end{aligned}$$

therefore $\boldsymbol{w}$ is orthogonal to both those matrices, hence we can take the cross product, which we already know

$$\boldsymbol{w} = k \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix} \qquad k \in \mathbb{R}$$

We can then conclude the following;

$$\boldsymbol{w} \in \ker(\boldsymbol{A}^\top) \Leftrightarrow \boldsymbol{w} \cdot \begin{bmatrix}1\\1\\1\end{bmatrix} = 0 = \boldsymbol{w} \cdot \begin{bmatrix}1\\0\\-1\end{bmatrix} \Leftrightarrow \forall \boldsymbol{y} \in \mathrm{Im}(\boldsymbol{A})\ [\boldsymbol{w} \cdot \boldsymbol{y} = 0]$$

This gives the result that $\boldsymbol{w}$ is in the null space of $\boldsymbol{A}^\top$ iff it is orthogonal to every vector in the image space of $\boldsymbol{A}$.

Our conjecture is the following; $\ker(\boldsymbol{A}^\top) \perp \mathrm{Im}(\boldsymbol{A})$, which is the same as saying

$$\boldsymbol{v} \in \ker(\boldsymbol{A}^\top) \Leftrightarrow \forall \boldsymbol{y} \in \mathrm{Im}(\boldsymbol{A})\ [\boldsymbol{y} \cdot \boldsymbol{v} = 0]$$

This can be proven generally as follows;

$$\begin{aligned}
&\boldsymbol{v} \in \ker(\boldsymbol{A}^\top) \\
\Leftrightarrow\ &\boldsymbol{A}^\top \boldsymbol{v} = \boldsymbol{0} && \text{by definition of the null space / kernel} \\
\Leftrightarrow\ &\forall i \in [1, n]\ [\boldsymbol{a_i} \cdot \boldsymbol{v} = 0] && (1) \\
\Leftrightarrow\ &\forall \boldsymbol{x} \in \mathbb{R}^n\ \left[\left(\sum_{i=1}^{n} x_i \boldsymbol{a_i}\right) \cdot \boldsymbol{v} = 0\right] && \text{generalise to any linear combination} \\
\Leftrightarrow\ &\boldsymbol{y} \in \mathrm{Im}(\boldsymbol{A})\ [\boldsymbol{y} \cdot \boldsymbol{v} = 0] && \blacksquare
\end{aligned}$$

(1) row i of $\boldsymbol{A}^\top \in \mathbb{R}^{n \times m}$ is the same as column i of $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, transposed, and all of the components must become 0

We say that $\mathbb{R}^m$ is the direct sum of $\ker(\boldsymbol{A}^\top)$ and $\mathrm{Im}(\boldsymbol{A})$. We can get the following result, following what was proven earlier;

$$\boldsymbol{v} \in \mathrm{Im}(\boldsymbol{A}) \wedge \boldsymbol{v} \in \ker(\boldsymbol{A}^\top) \Leftrightarrow \boldsymbol{v} \cdot \boldsymbol{v} = 0 \Leftrightarrow \boldsymbol{v} = \boldsymbol{0}$$

We want to prove the result for any vector $\boldsymbol{x} \in \mathbb{R}^m$, there exists a unique vector $\boldsymbol{x_r} \in \mathrm{Im}(\boldsymbol{A})$ and $\boldsymbol{x_n} \in \ker(\boldsymbol{A}^\top)$, such that $\boldsymbol{x} = \boldsymbol{x_r} + \boldsymbol{x_n}$, or formally;

$$\forall \boldsymbol{x} \in \mathbb{R}^m\ \exists! \boldsymbol{x_r} \in \mathrm{Im}(\boldsymbol{A}), \boldsymbol{x_n} \in \ker(\boldsymbol{A}^\top)\ [\boldsymbol{x} = \boldsymbol{x_r} + \boldsymbol{x_n}]$$

By taking the union of a basis of $\mathrm{Im}(\boldsymbol{A})$, let it be $(\boldsymbol{v_1}, \ldots, \boldsymbol{v_r})$, and a basis of $\ker(\boldsymbol{A}^\top)$, let it be $(\boldsymbol{v_{r+1}}, \ldots, \boldsymbol{v_m})$, we get a basis of $\mathbb{R}^m$: $(\boldsymbol{v_1}, \ldots, \boldsymbol{v_r}, \boldsymbol{v_{r+1}}, \ldots, \boldsymbol{v_m})$. This gives us the following result;

$$\boldsymbol{x} \in \mathbb{R}^m = \sum_{i=1}^{m} x_i \boldsymbol{v_i} = \underbrace{\sum_{i=1}^{r} x_i \boldsymbol{v_i}}_{\in \mathrm{Im}(\boldsymbol{A})} + \underbrace{\sum_{i=r+1}^{m} x_i \boldsymbol{v_i}}_{\in \ker(\boldsymbol{A}^\top)} = \boldsymbol{x_r} + \boldsymbol{x_n}$$

In order to prove uniqueness, we assume $\boldsymbol{x} = \boldsymbol{x_r} + \boldsymbol{x_n} = \boldsymbol{x'_r} + \boldsymbol{x'_n}$. The violet part only holds iff $\boldsymbol{x_r} - \boldsymbol{x'_r} = \boldsymbol{x'_n} - \boldsymbol{x_n}$. As they are both in their respective subspaces of $\mathrm{Im}(\boldsymbol{A})$ and $\ker(\boldsymbol{A}^\top)$, they must both be $\boldsymbol{0}$, as shown above. This gives the result $\boldsymbol{x_r} = \boldsymbol{x'_r} \wedge \boldsymbol{x_n} = \boldsymbol{x'_n}$, hence proving uniqueness.

5th February 2020

Positive Definiteness

The conditions for positive definiteness ($\forall \boldsymbol{x} \neq \boldsymbol{0}\ [\boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x} > 0]$) of $\boldsymbol{A}$ are as follows ;

- all diagonal elements of $\boldsymbol{A}$ must be > 0

  Take $\boldsymbol{x} = \boldsymbol{e_i}$, which is all 0, other than the i^{th} entry being 1. Then $\boldsymbol{e_i}^\top \boldsymbol{A} \boldsymbol{e_i} = a_{i,i}$, which has to be greater than 0.

- all principal minors of $\boldsymbol{A}$ must be positive definite.

  For $1 \leq k \leq n$, $\boldsymbol{A^{(k)}} \in \mathbb{R}^{k \times k}$ denotes the top left square submatrix of $\boldsymbol{A}$.

$$\begin{aligned} \boldsymbol{x} &= \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{0} \end{bmatrix} && \text{where } \boldsymbol{y} \in \mathbb{R}^k, \boldsymbol{0} \in \mathbb{R}^{n-k} \\ \boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x} &= \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{0} \end{bmatrix}^\top \boldsymbol{A} \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{0} \end{bmatrix} \\ &= \boldsymbol{y}^\top \boldsymbol{A^{(k)}} \boldsymbol{y} \end{aligned}$$

  So for any $\boldsymbol{y} \neq \boldsymbol{0}$, $\boldsymbol{x} \neq \boldsymbol{0}$, thus $\boldsymbol{y}^\top \boldsymbol{A^{(k)}} \boldsymbol{y} > 0 \Leftrightarrow \boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x} > 0$

- $|a_{i,j}| < \max\{a_{i,i}, a_{j,j}\}$

  For any entry not on the diagonal, there are two corresponding entries on the diagonal (one corresponding to the row, the other to the column);

$$\boldsymbol{A} = \begin{bmatrix} \ddots & & & & \\ & a_{i,i} & \cdots & a_{i,j} & \\ & & \ddots & \vdots & \\ & & & a_{j,j} & \\ & & & & \ddots \end{bmatrix}$$

$$\begin{aligned} &\text{take } \boldsymbol{x} = \boldsymbol{e_i} + \boldsymbol{e_j} \\ &\boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x} = (\boldsymbol{e_i} \pm \boldsymbol{e_j})^\top \boldsymbol{A} (\boldsymbol{e_i} \pm \boldsymbol{e_j}) \\ &\qquad = \left(\boldsymbol{e_i}^\top \pm \boldsymbol{e_j}^\top\right) \boldsymbol{A} (\boldsymbol{e_i} \pm \boldsymbol{e_j}) \\ &\qquad = \boldsymbol{e_i}^\top \boldsymbol{A} \boldsymbol{e_i} + \boldsymbol{e_j}^\top \boldsymbol{A} \boldsymbol{e_j} \pm \boldsymbol{e_i}^\top \boldsymbol{A} \boldsymbol{e_j} \pm \boldsymbol{e_j}^\top \boldsymbol{A} \boldsymbol{e_i} \\ &\qquad = a_{i,i} + a_{j,j} \pm 2a_{i,j} && \boldsymbol{A} \text{ is symmetric} \\ &\qquad > 0 && \Rightarrow \\ &-a_{i,j} < \frac{a_{i,i} + a_{j,j}}{2} && \text{diagonal terms are positive} \\ &\text{or } a_{i,j} < \frac{a_{i,i} + a_{j,j}}{2} && \Rightarrow \\ &|a_{i,j}| < \max\{a_{i,i}, a_{j,j}\} \end{aligned}$$

- all positive eigenvalues

For an example, take the following;

$$\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 3 \\ -1 & 10 & -1 \\ -3 & -1 & 2 \end{bmatrix}$$

However, we see that $|-3| = 3 > 2 = \max\{1, 2\}$, hence $\boldsymbol{A}$ is not positive definite

Least Square Method

For a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, where n is commonly the dimension of the data, and m the number of samples (such that each row is an entry), we can write $\boldsymbol{A}$ as;

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a^1} \\ \boldsymbol{a^2} \\ \vdots \\ \boldsymbol{a^m} \end{bmatrix} \text{ (all rows)}$$

The goal of the LSM is to find a $\boldsymbol{x} \in \mathbb{R}^n$, such that $\|\boldsymbol{Ax} - \boldsymbol{b}\|_2$ is minimised (which is equivalent to minimising $\|\boldsymbol{Ax} - \boldsymbol{b}\|_2^2$, as it is positive). From previous results, we can take $\boldsymbol{b} = \boldsymbol{b_r} + \boldsymbol{b_n}$, therefore we can minimise the following;

$$\begin{aligned}
\|\boldsymbol{Ax} - \boldsymbol{b_r} - \boldsymbol{b_n}\|_2^2 &= ((\boldsymbol{Ax} - \boldsymbol{b_r}) - \boldsymbol{b_n}) \cdot ((\boldsymbol{Ax} - \boldsymbol{b_r}) - \boldsymbol{b_n}) \\
&= (\boldsymbol{Ax} - \boldsymbol{b_r}) \cdot (\boldsymbol{Ax} - \boldsymbol{b_r}) + \boldsymbol{b_n} \cdot \boldsymbol{b_n} - 2\overbrace{(\boldsymbol{Ax} - \boldsymbol{b_r})}^{\in \mathrm{Im}(\boldsymbol{A})} \cdot \underbrace{\boldsymbol{b_n}}_{\in \ker(\boldsymbol{A}^\top)} && (1) \\
&= (\boldsymbol{Ax} - \boldsymbol{b_r}) \cdot (\boldsymbol{Ax} - \boldsymbol{b_r}) + \boldsymbol{b_n} \cdot \boldsymbol{b_n} && (2)
\end{aligned}$$

(1) the dot product is bilinear in both variables, hence we can expand out

(2) range space of $\boldsymbol{A}$ is perpendicular to the null space of $\boldsymbol{A}^\top$, hence the dot product is 0

Therefore, we want to minimise $(\boldsymbol{Ax} - \boldsymbol{b_r}) \cdot (\boldsymbol{Ax} - \boldsymbol{b_r}) + \boldsymbol{b_n} \cdot \boldsymbol{b_n}$. However, we have no control over $\boldsymbol{b_n}$, therefore we want to solve for $\boldsymbol{Ax} = \boldsymbol{b_r}$. This will always have a solution, as $\boldsymbol{b_r} \in \mathrm{Im}(\boldsymbol{A})$, which means that such a $\boldsymbol{x}$ exists by definition of the range space. We then claim the following;

$$\boldsymbol{Ax} = \boldsymbol{b_r} \Leftrightarrow \boldsymbol{A}^\top \boldsymbol{Ax} = \boldsymbol{A}^\top \boldsymbol{b}$$

Therefore, if the left hand side of the double implication has a solution, which we know it does, then the right hand side must also have a solution. The proof for this is done in both directions;

proving "⇒"

$$\begin{aligned}
\text{suppose } \boldsymbol{Ax} &= \boldsymbol{b_r} && \Rightarrow \\
\boldsymbol{A}^\top \boldsymbol{Ax} &= \boldsymbol{A}^\top \boldsymbol{b_r} && \text{pre-multiply} \\
&= \boldsymbol{A}^\top (\boldsymbol{b_r} + \boldsymbol{b_n}) && \text{in null space, hence } \boldsymbol{A}^\top \boldsymbol{b_n} = \boldsymbol{0} \\
&= \boldsymbol{A}^\top \boldsymbol{b} && \blacksquare
\end{aligned}$$

proving "⇐"

$$\begin{aligned}
\text{suppose } \boldsymbol{A}^\top \boldsymbol{Ax} &= \boldsymbol{A}^\top \boldsymbol{b} && \Rightarrow \\
\boldsymbol{A}^\top \boldsymbol{Ax} - \boldsymbol{A}^\top (\boldsymbol{b_r} + \boldsymbol{b_n}) &= \boldsymbol{0} && \Rightarrow \\
\boldsymbol{A}^\top (\boldsymbol{Ax} - \boldsymbol{b_r} - \boldsymbol{b_n}) &= \boldsymbol{0} && \Rightarrow \\
\boldsymbol{Ax} - \boldsymbol{b_r} - \boldsymbol{b_n} &\in \ker(\boldsymbol{A}^\top) && \text{property of adding in subspace}, \Rightarrow \\
\boldsymbol{Ax} - \boldsymbol{b_r} &\in \ker(\boldsymbol{A}^\top) && \text{and} \\
\boldsymbol{Ax} - \boldsymbol{b_r} &\in \mathrm{Im}(\boldsymbol{A}^\top) && \text{adding two elements of the range space}, \Rightarrow \\
\boldsymbol{Ax} - \boldsymbol{b_r} &= \boldsymbol{0} && \Rightarrow \\
\boldsymbol{Ax} &= \boldsymbol{b_r} && \blacksquare
\end{aligned}$$

An example of using this is as follows (note that $\boldsymbol{Ax} = \boldsymbol{b}$ is inconsistent);

$$\boldsymbol{A} = \begin{bmatrix} 2 & 2 \\ 1 & 2 \\ 2 & 0 \end{bmatrix}$$

$$\boldsymbol{b} = \begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix}$$

find $\boldsymbol{x}$ such that $\boldsymbol{Ax} = \boldsymbol{b}$,

$$\boldsymbol{A}^\top \boldsymbol{Ax} = \boldsymbol{A}^\top \boldsymbol{b} \qquad \Rightarrow$$

$$\begin{bmatrix} 9 & 6 \\ 6 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 2 \\ 2 & 2 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} 3 \\ 10 \end{bmatrix}$$

we know this is positive definite, and therefore there is a solution

$$x_1 = -1$$
$$x_2 = 2$$

6th February 2020

Using Cholesky Decomposition for LSM

While it's easy in some cases to solve with an inverse, we have a symmetric positive definite matrix in $\boldsymbol{A}^\top \boldsymbol{A}$. Therefore we have the following steps;

- goal is to solve $\boldsymbol{A}^\top \boldsymbol{Ax} = \boldsymbol{A}^\top \boldsymbol{b}$
- let $\boldsymbol{A}^\top \boldsymbol{A} = \boldsymbol{LL}^\top$, therefore we have $\boldsymbol{LL}^\top \boldsymbol{x} = \boldsymbol{A}^\top \boldsymbol{b}$
- let $\boldsymbol{y} = \boldsymbol{L}^\top \boldsymbol{x}$
- find $\boldsymbol{y}$ with forward substitution in $\boldsymbol{Ly} = \boldsymbol{A}^\top \boldsymbol{b}$
- find $\boldsymbol{x}$ with backward substitution in $\boldsymbol{y} = \boldsymbol{L}^\top \boldsymbol{x}$

First Application (Linear Regression) + Example

A simple application for this is to consider m samples of pairs (a_i, b_i). We want to find an affine line $y = s_1 x + s_0$, such that;

$$\sum_{i=1}^{m} e_i^2 \text{ is minimised, where the error term } e_i = |(s_1 a_i + s_0) - b_i|$$

In this case, we only have two unknowns, and therefore we can write the following;

$$\boldsymbol{A} = \begin{bmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \vdots \\ 1 & a_m \end{bmatrix}$$

$$\boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

$$\boldsymbol{z} = \begin{bmatrix} s_0 \\ s_1 \end{bmatrix}$$

$$\boldsymbol{Az} = \boldsymbol{b} \qquad \Rightarrow$$

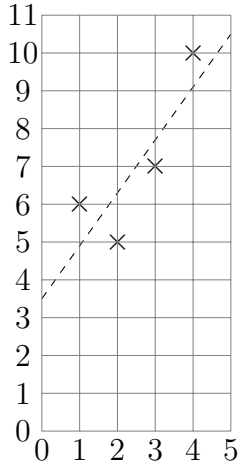
$$\begin{bmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \vdots \\ 1 & a_m \end{bmatrix} \begin{bmatrix} s_0 \\ s_1 \end{bmatrix} = \begin{bmatrix} s_0 + a_1 s_1 \\ s_0 + a_2 s_1 \\ \vdots \\ s_0 + a_m s_1 \end{bmatrix}$$
$$= \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

However, being able to solve it exactly is unlikely, especially for large datasets. In a real example, we have the following;

$$(a_1, b_1) = (1, 6), (2, 5), (3, 7), (4, 10)$$
$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$$
$$\boldsymbol{b} = \begin{bmatrix} 6 \\ 5 \\ 7 \\ 10 \end{bmatrix}$$
$$\boldsymbol{A}^\top \boldsymbol{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$$
$$= \begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix}$$
$$\boldsymbol{A}^\top \boldsymbol{b} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 6 \\ 5 \\ 7 \\ 10 \end{bmatrix}$$
$$= \begin{bmatrix} 28 \\ 77 \end{bmatrix}$$

by doing Gaussian elimination, we get the following results

$$s_0 = \frac{7}{2}$$
$$s_1 = \frac{7}{5}$$



Higher Dimensions

In order to generalise this to higher dimensions, for each (a_i, b_i), we now have a vector $\boldsymbol{a_i} \in \mathbb{R}^n$, and a scalar $b_i \in \mathbb{R}$, for $i = 1, \ldots, m$. The affine line is now no longer as simple, and is therefore written as;

$$y = \boldsymbol{x}^\top \boldsymbol{s} + s_0 = \boldsymbol{x} \cdot \boldsymbol{s} + s_0 = s_0 + \sum_{j=1}^{n} x_j s_j, \text{ where } \boldsymbol{s} \in \mathbb{R}^n, \text{ and } s_0 \in \mathbb{R}.$$

Setting this out, we have the following;

$$\boldsymbol{a_i} = \begin{bmatrix} a_{i,1} \\ \vdots \\ a_{i,n} \end{bmatrix} \qquad \text{for } i = 1, \ldots, m$$

$$\boldsymbol{A} = \begin{bmatrix} 1 & a_{1,1} & \cdots & a_{1,n} \\ 1 & a_{2,1} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_{2,1} & \cdots & a_{m,n} \end{bmatrix}$$

$$\boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

$$\boldsymbol{z} = \begin{bmatrix} s_0 \\ \boldsymbol{s} \end{bmatrix} \qquad \boldsymbol{s} \in \mathbb{R}^n$$

Second Application (Polynomial Regression)

Another application of this is to suppose data is (y_i, t_i), where $y_i \in \mathbb{R}$ and $t_i \in \mathbb{R}$ (considered as "time" - but is any continuous variable). We hypothesise the following;

$$y(t) = \sum_{j=1}^{n} P_j f_j(t)$$

where P_js are the parameters of the model, and f_js are basic functions, for **example**; $f_1(t) = 1, f_2(t) = t, \ldots, f_n(t) = t^{n-1}$. This is called **polynomial regression**. We know (we're given) f_js, but we want to find the P_js.

$$\begin{aligned} \boldsymbol{Ax} &= \boldsymbol{y} \\ \boldsymbol{A} &= (a_{i,j}) \\ a_{i,j} &= f_j(t_i) \\ \boldsymbol{x} &= \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_m \end{bmatrix} \end{aligned}$$

An example of this is to determine height (h), and the coefficient of gravity (g). We hypothesise that the distance with respect to time, away from the start, is

$$y(t) = h - g\frac{t^2}{2}$$

We assume that h and g are the unknown parameters, and define the basic functions as follows (note that the pairs (y_i, t_i) can be obtained with experiments);

$$f_1(t) = 1$$
$$f_2(t) = -\frac{t^2}{2}$$
$$\boldsymbol{A} = \begin{bmatrix} 1 & -\frac{t_1^2}{2} \\ \vdots & \vdots \\ 1 & -\frac{t_m^2}{2} \end{bmatrix}$$
$$\boldsymbol{x} = \begin{bmatrix} h \\ g \end{bmatrix}$$
$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

This can now be solved with the standard LSM.

QR Decomposition

We define $\boldsymbol{Q} \in \mathbb{R}^{m \times n}$ to be an orthogonal matrix if $\boldsymbol{Q} = [\boldsymbol{q_1}, \ldots, \boldsymbol{q_n}]$, the column vectors are unit vectors with respect to the ℓ_2-norm, and are pairwise orthogonal. This is only possible if $n \leq m$. It's important to note that $\boldsymbol{Q}^\top \boldsymbol{Q} = \boldsymbol{I}_n$, and $\boldsymbol{Q}\boldsymbol{Q}^\top = \boldsymbol{I}_m$.

The goal is to find $\boldsymbol{Q}$ and $\boldsymbol{R}$ such that $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$, where $\boldsymbol{Q}$ is an orthogonal matrix, and $\boldsymbol{R}$ is an upper triangular matrix. It's trivial to see that $\boldsymbol{R} = \boldsymbol{Q}^\top \boldsymbol{A}$, due to the previously mentioned property. This is used to easily solve the following (since $\boldsymbol{R}$ is upper triangular, we can employ backward substitution);

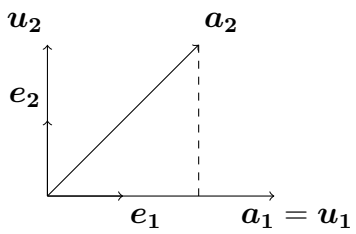
$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \Rightarrow \boldsymbol{Q}\boldsymbol{R}\boldsymbol{x} = \boldsymbol{b} \Rightarrow \boldsymbol{R}\boldsymbol{x} = \boldsymbol{Q}^\top \boldsymbol{b}$$

Gram-Schmidt Process

Suppose we have a set of linearly independent vectors $\boldsymbol{a_1}, \ldots, \boldsymbol{a_n} \in \mathbb{R}^m$. The Gram-Schmidt process is as follows;

$$\begin{aligned}
\boldsymbol{u_1} &= \boldsymbol{a_1} & \boldsymbol{e_1} &= \frac{\boldsymbol{u_1}}{\|\boldsymbol{u_1}\|} \\
\boldsymbol{u_2} &= \boldsymbol{a_2} - (\boldsymbol{e_1} \cdot \boldsymbol{a_2})\boldsymbol{e_1} & \boldsymbol{e_2} &= \frac{\boldsymbol{u_2}}{\|\boldsymbol{u_2}\|} \\
\boldsymbol{u_3} &= \boldsymbol{a_3} - (\boldsymbol{e_1} \cdot \boldsymbol{a_3})\boldsymbol{e_1} - (\boldsymbol{e_2} \cdot \boldsymbol{a_3})\boldsymbol{e_2} & \boldsymbol{e_2} &= \frac{\boldsymbol{u_3}}{\|\boldsymbol{u_3}\|} \\
&\vdots & &\vdots \\
\boldsymbol{u_n} &= \boldsymbol{a_n} - \sum_{j=1}^{n-1} (\boldsymbol{e_j} \cdot \boldsymbol{a_n})\boldsymbol{e_j} & \boldsymbol{e_n} &= \frac{\boldsymbol{u_n}}{\|\boldsymbol{u_n}\|}
\end{aligned}$$

Geometrically, we can visualise it as the following;



In general, for $1 \leq j \leq n$, we have

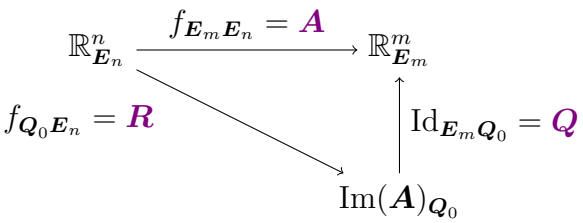
$$\boldsymbol{a_j} = (\boldsymbol{e_1} \cdot \boldsymbol{a_j})\boldsymbol{e_1} + (\boldsymbol{e_2} \cdot \boldsymbol{a_j})\boldsymbol{e_2} + \cdots + (\boldsymbol{e_j} \cdot \boldsymbol{a_j})\boldsymbol{e_j}$$

First Method for QR Decomposition (Gram-Schmidt Process)

Take a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, where $\boldsymbol{A} = [\boldsymbol{a_1}, \ldots, \boldsymbol{a_n}]$, assuming that $\boldsymbol{a_1}, \ldots, \boldsymbol{a_n}$ are linearly independent. We then apply Gram-Schmidt to $\boldsymbol{a_1}, \ldots, \boldsymbol{a_n}$ as described above to get $\boldsymbol{Q}_0 = (\boldsymbol{e_1}, \ldots, \boldsymbol{e_n})$, where. Note that for $1 \leq j \leq n$, $\boldsymbol{e_j} \cdot \boldsymbol{e_j} = 1$, and for $j \neq k$, $\boldsymbol{e_j} \cdot \boldsymbol{e_k} = 0$. Let $\boldsymbol{Q} = [\boldsymbol{e_1}, \ldots, \boldsymbol{e_n}]$, which is orthogonal. While we can see that $\boldsymbol{R} = \boldsymbol{Q}^\top \boldsymbol{A}$, we can write this as;

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{e_1} \cdot \boldsymbol{a_1} & \boldsymbol{e_1} \cdot \boldsymbol{a_2} & \boldsymbol{e_1} \cdot \boldsymbol{a_3} & \cdots & \boldsymbol{e_1} \cdot \boldsymbol{a_n} \\ & \boldsymbol{e_2} \cdot \boldsymbol{a_2} & \boldsymbol{e_2} \cdot \boldsymbol{a_3} & \cdots & \boldsymbol{e_2} \cdot \boldsymbol{a_n} \\ & & \boldsymbol{e_3} \cdot \boldsymbol{a_3} & \cdots & \boldsymbol{e_3} \cdot \boldsymbol{a_n} \\ & & & \ddots & \vdots \\ & & & & \boldsymbol{e_n} \cdot \boldsymbol{a_n} \end{bmatrix} \in \mathbb{R}^{n \times n}$$

We claim that the diagram below is valid when $\boldsymbol{A}$ represents the linear map $f : \mathbb{R}^n \to \mathbb{R}^m$, in the standard basis.



12th February 2020

Proof of Diagram

Continuing on from last week's lecture, we aim to prove two claims in the diagram;

- $f_{\boldsymbol{Q}_0 \boldsymbol{E}_n} = \boldsymbol{R}$

  The matrix $\boldsymbol{R}$ represents the linear map f from $\mathbb{R}^n$ (in the standard basis) to $\mathbb{R}^m$ in the new basis $(\boldsymbol{Q}_0)$.

  To begin, we observe how the standard basis is mapped under $\boldsymbol{A}$;

$$j^{\text{th}} \to \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \xmapsto{\boldsymbol{A} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \boldsymbol{a_j}} \boldsymbol{a_j} = (\boldsymbol{e_1} \cdot \boldsymbol{a_j})\boldsymbol{e_1} + (\boldsymbol{e_2} \cdot \boldsymbol{a_j})\boldsymbol{e_2} + \cdots + (\boldsymbol{e_j} \cdot \boldsymbol{a_j})\boldsymbol{e_j} + 0\boldsymbol{e_{j+1}} + \cdots + 0\boldsymbol{e_n}$$

  Looking at these coefficients with respect to the basis formed by $(\boldsymbol{e_1}, \ldots, \boldsymbol{e_n})$, which are highlighted in violet, it's clear that it forms the j^{th} column of $\boldsymbol{R}$.

- $\mathrm{Id}_{\boldsymbol{E}_m \boldsymbol{Q}_0} = \boldsymbol{Q}$

  This is a change of basis in the image space of $\boldsymbol{A}$.

  This is trivial to show, as we want to look at what happens to the j^{th} vector in the new basis $\boldsymbol{Q}_0$. To write each of these matrices in the new (standard) basis, we simply do the following;

$$\boldsymbol{e_j} = e_{1,j}\begin{bmatrix}1\\0\\0\\\vdots\\0\end{bmatrix} + e_{2,j}\begin{bmatrix}0\\1\\0\\\vdots\\0\end{bmatrix} + \cdots + e_{m,j}\begin{bmatrix}0\\0\\\vdots\\0\\1\end{bmatrix}$$

Trivially, we get $\boldsymbol{Q} = [\boldsymbol{e_1}, \ldots, \boldsymbol{e_n}]$, as expected.

Note on Basis Change

To find the matrix representing $f_{\boldsymbol{DB}}$, which takes a vector in basis $\boldsymbol{B}$ to a vector in basis $\boldsymbol{D}$, we look at what happens to each basis vector.

$$\boldsymbol{b_j} \mapsto f(\boldsymbol{b_j}) = \delta_1\boldsymbol{d_1} + \delta_2\boldsymbol{d_1} + \cdots + \delta_m\boldsymbol{d_m}$$

Then the j^{th} column of the matrix is given as;

$$(f_{\boldsymbol{DB}})_j = \begin{bmatrix}\delta_1\\\delta_2\\\vdots\\\delta_m\end{bmatrix}$$

Householder Operator

I can't draw the geometric intuition nicely, at all, look at Panopto timestamp 36:37. However, imagine there is a plane, and there is some point $\boldsymbol{x}$ that exists. That point is orthogonally projected onto the plane, with the orthogonal projection matrix $\boldsymbol{P}$, such that the vector between the point $\boldsymbol{x}$ and $\boldsymbol{Px}$ is perpendicular to the plane. Now, take the perpendicular unit vector for the plane as $\boldsymbol{u}$, therefore the length of $\boldsymbol{x}$ to $\boldsymbol{Px}$ is $\boldsymbol{u}\cdot\boldsymbol{x}$. Imagine the plane as a mirror, and the image of $\boldsymbol{x}$ is on the other side of the plane, let that point be $H(\boldsymbol{x}) = \boldsymbol{Hx}$. Therefore, we can work out $\boldsymbol{H}$ as follows;

$$H(\boldsymbol{x}) = \boldsymbol{x} - 2(\boldsymbol{u}\cdot\boldsymbol{x})\boldsymbol{u} = \boldsymbol{x} - 2\boldsymbol{x}\boldsymbol{u}\boldsymbol{u}^\top = \underbrace{\left(\boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^\top\right)}_{\boldsymbol{H}}\boldsymbol{x}$$

We expect $\boldsymbol{H}^2$ to give us the identity, as if "reflect" a reflected point, it should give the original;

$$\begin{aligned}\boldsymbol{H}^2 &= \left(\boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^\top\right)\\ &= \boldsymbol{I} - 4\boldsymbol{u}\boldsymbol{u}^\top + 4\boldsymbol{u}\left(\boldsymbol{u}^\top\boldsymbol{u}\right)\boldsymbol{u}^\top\\ &= \boldsymbol{I} - 4\boldsymbol{u}\boldsymbol{u}^\top + 4\boldsymbol{u}\boldsymbol{u}^\top && \ell_2\text{-norm of unit vector is 1}\\ &= \boldsymbol{I} && \blacksquare\end{aligned}$$

Note that there are two distinct eigenvalues of $\boldsymbol{H}$; $\lambda_1 = 1$ has multiplicity $m_1 = m - 1$, where m is the dimension of the space. This is for each of the vectors $\boldsymbol{v_i}$ that span the plane, as they are mapped to themselves ($\boldsymbol{Hv_i} = H(\boldsymbol{v_i}) = \boldsymbol{v_i}$). There is also an additional eigenvalue $\lambda_2 = -1$, for $\boldsymbol{u}$, as we can see that $\boldsymbol{Hu} = H(\boldsymbol{u}) = -\boldsymbol{u}$.

Second Method for QR Decomposition (Householder Operator)

Take the matrix $\boldsymbol{A} = [\boldsymbol{a_1}, \ldots, \boldsymbol{a_n}]$. The Householder operator must preserve the length, in the following mapping;

$$H\left(\begin{bmatrix}a_{1,1}\\a_{2,1}\\\vdots\\a_{m,1}\end{bmatrix}\right) = \begin{bmatrix}\lambda\\0\\\vdots\\0\end{bmatrix}$$

$$\pm\lambda = \|\boldsymbol{a_1}\|$$

We want to send the vector $\boldsymbol{a_1}$ to the vector $\|\boldsymbol{a_1}\|\boldsymbol{e_1}$, where $\boldsymbol{e_1}$ is all 0, other than the first element.

$$\begin{aligned}
\boldsymbol{u} &= \frac{\boldsymbol{a_1} - \|\boldsymbol{a_1}\|\boldsymbol{e_1}}{\|\boldsymbol{a_1} - \|\boldsymbol{a_1}\|\boldsymbol{e_1}\|} \\
\boldsymbol{H_1} &= \boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^\top \\
\boldsymbol{H_1 A} &= \boldsymbol{H_1}[\boldsymbol{a_1}, \ldots, \boldsymbol{a_n}] \\
&= [\boldsymbol{H_1 a_1}, \ldots, \boldsymbol{H_1 a_n}] \\
&= \begin{bmatrix} \|\boldsymbol{a_1}\| & \times \\ \boldsymbol{0} & \boldsymbol{A_2} \end{bmatrix} \qquad \times \text{ is an unknown row, } \boldsymbol{A_2} \text{ is the bottom right submatrix}
\end{aligned}$$

Applying $\boldsymbol{H}$ to $\boldsymbol{A}$ gets rid of all entries in the first column, other than the first element. This is an iterative process, where every iteration brings us closer to an upper triangular matrix. Note that the Householder matrix, $\boldsymbol{H_2}$, applied to $\boldsymbol{A_2}$ is 1 smaller on each dimension. To account for this, a top row, and left column is added, of all 0s, other than the top left element, which is 1.

$$\boldsymbol{H_n} \cdots \boldsymbol{H_1 A} = \boldsymbol{R}$$

By reversing this, we can get the following result;

$$\boldsymbol{A} = \boldsymbol{H_1} \cdots \boldsymbol{H_n R}$$

Using the Householder operator to do QR decomposition isn't examinable.

13th February 2020

In the second half of the course, we cover;

- condition number
- convergence and fixed point problems
- iterative solution of linear equations and eigenvectors
- Laplace and Fourier transforms
- functions of several variables and optimisation

Condition

When a matrix is very close to singular, the results can be unreliable when computed by a computer, as the results can push up to the precision of the machine. Consider the following set of linear equations;

$$\begin{aligned}
x + y &= 1 \\
x + \alpha y &= 0 \qquad \Rightarrow \\
x &= 1 - \frac{1}{1-\alpha} \\
y &= \frac{1}{1-\alpha}
\end{aligned}$$

The solutions are inconsistent if $\alpha = 1$, and will become very large (and likely inaccurate on a computer) when $\alpha \sim 1$.

The condition number of a problem P is the maximum size-ratio, where d_1 and d_2 are alternate inputs, and $s(d_1)$ and $s(d_2)$ are the corresponding solutions;

$$\kappa(P) = \max_{d_1, d_2} \frac{\|s(d_1) - s(d_2)\|}{\|d_1 - d_2\|}$$

The key point is we are looking for the worst case.

Recap on Norms

We want to recall that for each vector norm $(1, 2, \infty)$, the subordinate (consistent or compatible) matrix norm is

$$\|\boldsymbol{A}\| = \max\left\{\frac{\|\boldsymbol{Ax}\|}{\|\boldsymbol{x}\|} : \boldsymbol{x} \neq \boldsymbol{0}\right\} = \max\{\|\boldsymbol{Au}\| : \|\boldsymbol{u}\| = 1\}$$

This means that we have a lower bound of the matrix norm;

$$\forall \boldsymbol{x} \neq \boldsymbol{0} \ \left[\|\boldsymbol{A}\| \geq \frac{\|\boldsymbol{Ax}\|}{\|\boldsymbol{x}\|}\right]$$

A brief recap on the terms;

- consistent norm $\|\boldsymbol{Ax}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{x}\|$
- compatible norm consistent norm defined on square matrix
- all subordinate (also called induced) norms are consistent

Perturbation

Consider the system of linear equations $\boldsymbol{Ax} = \boldsymbol{b}$. Make a small change to $\boldsymbol{b}$, such that we have $\boldsymbol{b} + \delta\boldsymbol{b}$. Note that $\boldsymbol{x_b}$ is the change in $\boldsymbol{x}$ due to a change in $\boldsymbol{b}$. We then have the following;

$$\begin{aligned}
\boldsymbol{Ax} + \boldsymbol{A}\delta\boldsymbol{x_b} &= \boldsymbol{b} + \delta\boldsymbol{b} && \text{we know } \boldsymbol{Ax} = \boldsymbol{b}, \Rightarrow \\
\boldsymbol{A}\delta\boldsymbol{x_b} &= \delta\boldsymbol{b} && \text{assuming a solution exists}, \Rightarrow \\
\delta\boldsymbol{x_b} &= \boldsymbol{A}^{-1} && \text{by consistency}, \Rightarrow \\
\|\delta\boldsymbol{x_b}\| &\leq \|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{b}\| && \text{looking at \% increase}, \Rightarrow \\
\frac{\|\delta\boldsymbol{x_b}\|}{\|\boldsymbol{x}\|} &\leq \frac{\|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{b}\|}{\|\boldsymbol{x}\|} && (1)
\end{aligned}$$

we now briefly look at $\boldsymbol{Ax} = \boldsymbol{b}$

$$\begin{aligned}
\boldsymbol{Ax} &= \boldsymbol{b} && \Rightarrow \\
\|\boldsymbol{b}\| &\leq \|\boldsymbol{A}\|\|\boldsymbol{x}\| && \Rightarrow \\
\|\boldsymbol{x}\| &\geq \frac{\|\boldsymbol{b}\|}{\|\boldsymbol{A}\|}
\end{aligned}$$

we can replace the $\|\boldsymbol{x}\|$ in the RHS of (1) with something smaller, as it won't affect the inequality

$$\begin{aligned}
\frac{\|\delta\boldsymbol{x_b}\|}{\|\boldsymbol{x}\|} &\leq \frac{\|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{b}\|\|\boldsymbol{A}\|}{\|\boldsymbol{b}\|} && \Rightarrow \\
\frac{\|\delta\boldsymbol{x_b}\|}{\|\boldsymbol{x}\|} &\leq \|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|\frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}
\end{aligned}$$

Condition Number

Similarly, we can do the same for a small change in $\boldsymbol{A}$. As such, $\boldsymbol{x_A}$ is the change in $\boldsymbol{x}$ due to a change in $\boldsymbol{B}$.

$$\begin{aligned}
(\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x_A}) &= \boldsymbol{b} && \Rightarrow \\
\boldsymbol{Ax} + \boldsymbol{A}\delta\boldsymbol{x_A} + \delta\boldsymbol{A}\boldsymbol{x} + \delta\boldsymbol{A}\delta\boldsymbol{x_A} &= \boldsymbol{b} && (1), \Rightarrow \\
\boldsymbol{Ax} + \boldsymbol{A}\delta\boldsymbol{x_A} + \delta\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b} && \text{we know } \boldsymbol{Ax} = \boldsymbol{b}, \Rightarrow \\
\boldsymbol{A}\delta\boldsymbol{x_A} + \delta\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{0} && \Rightarrow \\
\boldsymbol{A}\delta\boldsymbol{x_A} &= -\delta\boldsymbol{A}\boldsymbol{x} && \text{assuming a solution exists}, \Rightarrow \\
\delta\boldsymbol{x_A} &= -\boldsymbol{A}^{-1}\delta\boldsymbol{A}\boldsymbol{x} && \text{by consistency, norm is positive}, \Rightarrow \\
\|\delta\boldsymbol{x_A}\| &\leq \|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{A}\|\|\boldsymbol{x}\| && \Rightarrow
\end{aligned}$$

$$\frac{\|\delta \boldsymbol{x_A}\|}{\|\boldsymbol{x}\|} \leq \|\boldsymbol{A}^{-1}\|\|\delta \boldsymbol{A}\| \qquad \text{we want relative change,} \Rightarrow$$
$$\frac{\|\delta \boldsymbol{x_A}\|}{\|\boldsymbol{x}\|} \leq \|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}$$

As such, we can define the condition of a matrix $\boldsymbol{A}$ as follows;

$$\text{cond}(\boldsymbol{A}) = \|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|$$

This gives us the following useful property, allowing the condition to be an upper bound on the worst case;

$$\text{cond}(\boldsymbol{A}) \geq \max\left\{ \frac{\frac{\|\delta \boldsymbol{x_b}\|}{\|\boldsymbol{x}\|}}{\frac{\|\delta \boldsymbol{b}\|}{\|\boldsymbol{b}\|}}, \frac{\frac{\|\delta \boldsymbol{x_A}\|}{\|\boldsymbol{x}\|}}{\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}} \right\}$$

Application to LSM

This is a problem that's well known to be numerically unstable. As usual, we take the case where there are m equations in n variables, where $m > n$. This gives us a non-square matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$. We also note that for the ℓ_2-norm, the condition of $\boldsymbol{A}^\top \boldsymbol{A}$ is the square of the condition of $\boldsymbol{A}$. Essentially, we are multiplying the norms of 4 matrices together, therefore we often get a large norm. We can work through the following example;

$$\begin{aligned}
\boldsymbol{A} &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 10^{-4} \end{bmatrix} \\
\boldsymbol{A}^\top \boldsymbol{A} &= \begin{bmatrix} 2 & 2 \\ 2 & 2 + 10^{-8} \end{bmatrix} \\
\left(\boldsymbol{A}^\top \boldsymbol{A}\right)^{-1} &= 0.5 \cdot 10^8 \cdot \begin{bmatrix} 2 + 10^{-8} & -2 \\ -2 & 2 \end{bmatrix} \\
\text{cond}\left(\boldsymbol{A}^\top \boldsymbol{A}\right)^{-1} &= (4 + 10^{-8})(4 + 10^{-8}) \cdot 0.5 \cdot 10^8 \\
&\approx 8 \cdot 10^8
\end{aligned}$$

While there is no threshold for deciding whether a matrix is ill-conditioned, the rule of thumb is that for some condition number $\kappa(\boldsymbol{A})$, you can lose $\log_{10}(\kappa(\boldsymbol{A}))$ significant figures in accuracy. For example, if we had a accurate starting point, for example something with 12 decimal places, and $\kappa = 1000000$, the precision will only reduce to 6 decimal places.

The residual vectors are only reliable indicators of the accuracy **if the problem is well-conditioned**;

$$\boldsymbol{r} = \boldsymbol{A}\boldsymbol{x_0} - \boldsymbol{b}$$

Metric Spaces

Take a non-empty set of points S, with a (distance) function $d : S \times S \to \mathbb{R}$ (the metric of the space), which satisfies ($\forall x, y, z \in S$);

(1) $d(x, x) = 0$

(2) $d(x, y) > 0$ given $x \neq y$

(3) $d(x, y) = d(y, x)$ (symmetry)

(4) $d(x, y) \leq d(x, z) + d(z, y)$ (triangle inequality)

This is true for distances in real spaces $\mathbb{R}, \mathbb{R}^2, \mathbb{R}^3, \ldots$ Another example of this is vector norms, as they satisfy the above properties.

We can justify that the discrete metric, defined below, satisfies those conditions;

$$d(x,y) = \begin{cases} 0 & x = y \\ 1 & x \neq y \end{cases}$$

Trivially, the first 3 conditions are satisfied by the definition, however the last takes more work to prove. We know that $d(x,y)$ is 0 or 1, and $d(x,z)+d(z,y)$ is 0, 1, or 2. If the latter is 1 or 2, there is no work to be done. However, consider the case where the former is 1, and the latter is 0. To start, we assume $d(x,y) = 1 \Rightarrow x \neq y$. We also assume $d(x,z) + d(z,y) = 0$, therefore $d(x,z) = 0$ and $d(z,y) = 0$. As such, we have $x = z$ and $z = y$, giving us $x = y$, which contradicts our initial assumption. Therefore that case **cannot** happen, and the last condition is satisfied.

Convergence

Briefly, we recap convergence from last year's **CO145**. To say that $a_1, a_2, \ldots$ converges to some limit $\ell \in \mathbb{R}$, which is written as;

$$a_n \to \ell \text{ as } n \to \infty$$
$$\text{or}$$
$$\lim_{n\to\infty} a_n = \ell$$

This is true iff (from first year notes)

$$\forall \epsilon > 0 \exists N \in \mathbb{N}\ [\forall n > N\ [|a_n - \ell| < \epsilon]]$$

Which holds iff (Cauchy)

$$\forall \epsilon > 0 \exists N \in \mathbb{N}\ [\forall n, m > N\ [|a_n - a_m| < \epsilon]]$$

The second method is useful as we don't need to know the limit (when it exists), such as in a recurrence relation or a recursively defined function, and can also be used as a test for divergence.

19th February 2020

Revised Matrix Norm Terminology

- a matrix norm $\|\cdot\|_{\alpha,\beta}$ is **consistent** relative to vector-norms α on $\mathbb{R}^m$ and β on $\mathbb{R}^n$ if

$$\forall \boldsymbol{A} \in \mathbb{R}^{m\times n}, \boldsymbol{x} \in \mathbb{R}^n\ [\|\boldsymbol{Ax}\|_\alpha \leq \|\boldsymbol{A}\|_{\alpha,\beta}\|\boldsymbol{x}\|_\beta]$$

- any induced / subordinate matrix norm is also sub-multiplicative because;

$$\|\boldsymbol{AB}\| = \max\{\|\boldsymbol{ABu}\| : \|\boldsymbol{u}\| = 1\} \leq \max\{\|\boldsymbol{A}\|\|\boldsymbol{B}\|\|\boldsymbol{u}\| : \|\boldsymbol{u}\| = 1\} = \|\boldsymbol{A}\|\|\boldsymbol{B}\|$$

Cauchy Sequence

Generalising limits to metric spaces (let it be S in lieu of repeating myself), we have the following definitions;

- a sequence $x_1, x_2, \cdots \in S$ converges to a limit $\ell \in S$ iff

$$\forall \epsilon > 0 \exists N \in \mathbb{N}\ [\forall n > N\ [d(x_n, \ell) < \epsilon]]$$

- a sequence $\{x_n\}$ is called a **Cauchy sequence** iff

$$\forall \epsilon > 0 \exists N \in \mathbb{N}\ [\forall n, m > N\ [d(x_n, x_m) < \epsilon]]$$

- a metric space in which every Cauchy sequence has a limit in S is **complete**

If a sequence $\{x_n\}$ converges in S, then $\{x_n\}$ is Cauchy. This means that any converging sequence is Cauchy, but not every Cauchy sequence is convergent. If every Cauchy sequence is convergent, then S is complete.

Application

We will mainly be working in real vector spaces $\mathbb{R}^k$. It can also be shown that $\mathbb{R}^k$ is complete for all $k \geq 1$ (proof not needed for this course / degree). Since these metric spaces are complete, we can then say that any sequence we encounter (in real vector spaces) is convergent iff it is Cauchy, which is relatively easy to test.

Uniqueness of Limits

This is an example of the type of reasoning we are expected to do in this course. We claim that a convergent sequence $\{x_n\} \in \mathbb{R}^k$ has a unique limit. The proof is as follows;

Suppose $x_n \to \ell_1$ and $x_n \to \ell_2$, as $n \to \infty$. By the last property of the metric, we can say (for all n);

$$d(\ell_1, \ell_2) \leq d(\ell_1, x_n) + d(x_n, \ell_2)$$

For a sufficiently large n, we can say the following;

$$d(\ell_1, \ell_2) < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon$$

Note that in the step above, we take any number, therefore $\frac{\epsilon}{2}$ is just as good as ϵ. Since ϵ is **any** positive number (by definition of a metric);

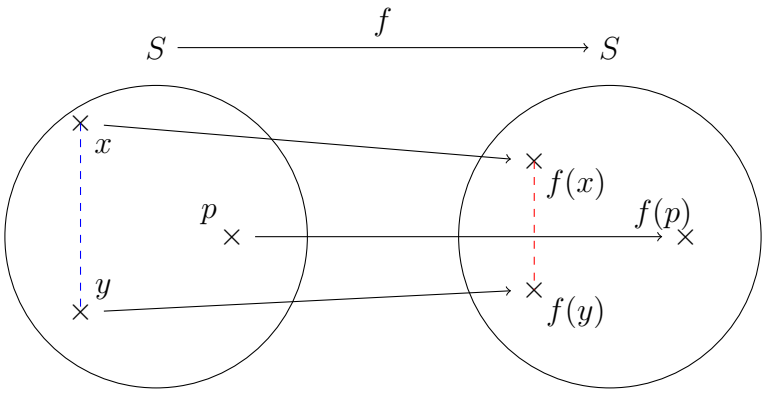
$$d(\ell_1, \ell_2) < \epsilon \Rightarrow d(\ell_1, \ell_2) = 0 \Leftrightarrow \ell_1 = \ell_2$$

Fixed Points

Let $f : S \to S$. A point $p \in S$ is a fixed point (or fixpoint) if $f(p) = p$. Additionally, a function f is a **contraction** of S if there exists a contraction constant $0 < \alpha < 1$;

$$\forall x, y \in S \; [d(f(x), f(y)) \leq \alpha d(x, y)]$$

Visually, we can represent it as the following;



Fixed Point Theorem

The fixed point theorem states that a continuous contraction f of a complete metric space S has a unique fixed point. The proof for this is as follows;

Take any point $x \in S$, we define $\{p_n\}$ as;

$$\begin{aligned}
p_0 &= x \\
p_{n+1} &= f(p_n) \\
d(p_{n+1}, p_n) &= d(f(p_n), f(p_{n-1})) \\
&\leq \alpha d(p_n, p_{n-1}) \\
&= \alpha d(f(p_{n-1}), f(p_{n-2})) \\
&\leq \alpha^2 d(p_{n-1}, p_{n-2}) \\
&\vdots \\
&\leq \alpha^n c \qquad\qquad \text{let } c = d(p_1, p_0)
\end{aligned}$$

We then assume $m > n$ (somewhat redundant by symmetry);

$$\begin{aligned}
d(p_m, p_n) &\leq d(p_m, p_{m-1}) + \cdots + d(p_{n+1}, p_n) \\
&= \sum_{k=n}^{m-1} d(p_{k+1}, p_k) \\
&\leq \sum_{k=n}^{m-1} \alpha^k c \\
&= c \sum_{k=n}^{m-1} \alpha^k && \text{notice the geometric series} \\
&= c\frac{\alpha^n - \alpha^m}{1-\alpha} && \alpha^m > 0 \\
&< \frac{c\alpha^n}{1-\alpha}
\end{aligned}$$

As $m, n \to \infty$, $\alpha^n \to 0$, so we have $d(p_m, p_n) \to 0$. This means that the points converge, therefore $\{p_n\}$ is Cauchy. With the completion property, we know that $p_n \to p$ for some point $p \in S$. To prove that it is a fixed point, we need to employ the continuity property of f;

$$\begin{aligned}
f(p) &= f(\lim_{n\to\infty} p_n) \\
&= \lim_{n\to\infty} f(p_n) && \text{by continuity, somehow?} \\
&= \lim_{n\to\infty} p_{n+1} \\
&= p
\end{aligned}$$

Finally, to prove the uniqueness of this, take two points $p, p' \in S$, both assumed to be fixpoints;

$$\begin{aligned}
d(p, p') &= d(f(p), f(p')) && \text{as they are fixpoints} \\
&\leq \alpha d(p, p') && \text{by contraction}
\end{aligned}$$

As $0 < \alpha < 1$, $d(p, p') = 0$, hence $p = p'$, therefore the fixpoint is unique.

Iterative Solutions of Linear Equations

Split a square matrix $\boldsymbol{A}$ to $\boldsymbol{M} = \boldsymbol{A} - \boldsymbol{N}$, where $\boldsymbol{M}$ is non-singular.

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b} && \Rightarrow \\
\boldsymbol{M}\boldsymbol{x} &= \boldsymbol{b} - \boldsymbol{N}\boldsymbol{x} && \Rightarrow \\
\boldsymbol{x} &= \boldsymbol{M}^{-1}\boldsymbol{b} - \boldsymbol{M}^{-1}\boldsymbol{N}\boldsymbol{x} && \Rightarrow \\
\boldsymbol{x} &= \boldsymbol{G}\boldsymbol{x} + \boldsymbol{c} && \text{where } \boldsymbol{G} = -\boldsymbol{M}^{-1}\boldsymbol{N}\boldsymbol{x}, \boldsymbol{c} = \boldsymbol{M}^{-1}\boldsymbol{b}
\end{aligned}$$

We then define the iteration to be as follows, with a starting point $\boldsymbol{x^{(0)}}$;

$$\boldsymbol{x^{(k+1)}} = \boldsymbol{G}\boldsymbol{x^{(k)}} + \boldsymbol{c}$$

20th February 2020

Proof of Convergence

We want to prove that for any consistent matrix norm, if $\|\boldsymbol{G}\| < 1$, then the iterative sequence converges for any starting $\boldsymbol{x^{(0)}}$.

Suppose $\boldsymbol{x}^{(k)} \to \hat{\boldsymbol{x}}$, where $\hat{\boldsymbol{x}} = \boldsymbol{G}\hat{\boldsymbol{x}} + \boldsymbol{c}$. We want to define $\boldsymbol{y}^{(k+1)}$ as follows;

$$\begin{aligned}
\boldsymbol{y}^{(k+1)} &= \boldsymbol{x}^{(k+1)} - \hat{\boldsymbol{x}} \\
&= \boldsymbol{G}\boldsymbol{x}^{(k)} + \boldsymbol{c} - (\boldsymbol{G}\hat{\boldsymbol{x}} + \boldsymbol{c}) \\
&= \boldsymbol{G}\boldsymbol{x}^{(k)} - \boldsymbol{G}\hat{\boldsymbol{x}} \\
&= \boldsymbol{G}(\boldsymbol{x}^{(k)} - \hat{\boldsymbol{x}}) \\
&= \boldsymbol{G}\boldsymbol{y}^{(k)}
\end{aligned}$$

From this, we can use the property that $\boldsymbol{G}$ is consistent as follows;

$$\begin{aligned}
\|\boldsymbol{y}^{(k+1)}\| &\leq \|\boldsymbol{G}\|\|\boldsymbol{y}^{(k)}\| \\
&\leq \|\boldsymbol{G}\|^2\|\boldsymbol{y}^{(k-1)}\| \\
&\leq \|\boldsymbol{G}\|^3\|\boldsymbol{y}^{(k-2)}\| \\
&\vdots \\
&\leq \|\boldsymbol{G}\|^{k+1}\|\boldsymbol{y}^{(0)}\|
\end{aligned}$$

However, we know that $\|\boldsymbol{G}\|^{k+1}\|\boldsymbol{y}^{(0)}\| \to 0$ when $k \to \infty$, as $\|\boldsymbol{G}\| < 1$. Therefore, $\boldsymbol{y}^{(k+1)} \to \boldsymbol{0}$ by the property of norms, hence $\boldsymbol{x}^{(k+1)} \to \hat{\boldsymbol{x}}$

Common Split

Assume, without loss of generality, that $\boldsymbol{A}$ has no zeroes on the diagonal. If it does, we can relabel the variables x_i to permute the rows and columns to have no zeroes on the diagonal.

We then define the following;

$$\boldsymbol{A} = \boldsymbol{D} + \boldsymbol{L} + \boldsymbol{U}$$

where $\boldsymbol{D}$ is the diagonal component, $\boldsymbol{L}$ is the **strict** lower triangular component (hence without the diagonal), and similar for $\boldsymbol{U}$.

Jacobi Method

In the Jacobi method, we set the matrices (and vector) as follows;

$$\begin{aligned}
\boldsymbol{M} &= \boldsymbol{D} \\
\boldsymbol{N} &= \boldsymbol{L} + \boldsymbol{U} \\
\boldsymbol{G} &= -\boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U}) \\
\boldsymbol{c} &= \boldsymbol{D}^{-1}\boldsymbol{b}
\end{aligned}$$

therefore $\boldsymbol{x}^{(k+1)} = \boldsymbol{G}\boldsymbol{x}^{(k)} + \boldsymbol{c}$ yields

$$x_i^{(k+1)} = \frac{1}{a_{i,i}}\left(b_i - \sum_{j \neq i} a_{j,j}x_j^{(k)}\right)$$

This is easily done in parallel.

Gauss-Seidel Method

$$\begin{aligned}
\boldsymbol{M} &= \boldsymbol{D} + \boldsymbol{L} \\
\boldsymbol{N} &= \boldsymbol{U} \\
\boldsymbol{G} &= -(\boldsymbol{D} + \boldsymbol{L})^{-1}\boldsymbol{U} \\
\boldsymbol{c} &= (\boldsymbol{D} + \boldsymbol{L})^{-1}\boldsymbol{b} \\
\text{therefore } \boldsymbol{x^{(k+1)}} &= \boldsymbol{G}\boldsymbol{x^{(k)}} + \boldsymbol{c} \qquad \text{yields}
\end{aligned}$$

$$x_i^{(k+1)} = \frac{1}{a_{i,i}}\left(b_i - \sum_{j=1}^{i-1} a_{i,j}x_j^{(k+1)} - \sum_{j=i+1}^{m} a_{i,j}x_j^{(k)}\right)$$

This gives faster convergence as we're able to use more up-to-date data. Working from the first column, we get;

$$x_1^{(k+1)} = \frac{1}{a_{1,1}}\left(b_1 - \sum_{j=2}^{m} a_{1,j}x_j^{(k)}\right)$$

$$x_2^{(k+1)} = \frac{1}{a_{2,2}}\left(b_2 - a_{2,1}x_1^{(k+1)} - \sum_{j=3}^{m} a_{2,j}x_j^{(k)}\right) \qquad \text{using the previous result}$$

$$\vdots$$

However, this must be done in order, and therefore cannot be done in parallel like in Jacobi.

Convergence Conditions

We define a square matrix $\boldsymbol{A}$ as **strictly** row diagonally dominant for the i^{th} row if

$$|a_{i,i}| > \sum_{j \neq i} |a_{i,j}|$$

This means that the sum of the moduli of the all the elements on the row (except the diagonal element) is less than the modulus of the diagonal element. The theorem is that this is a sufficient condition for both Jacobi and Gauss-Seidel to converge. Below, we will do the proof for Jacobi;

(1) show $\|\boldsymbol{G}\| < 1$, where $\boldsymbol{G} = -\boldsymbol{D}^{-1}\boldsymbol{N}$, using the consistent ℓ_∞-norm

(2) strict diagonal dominance implies all absolute row sums are less than 1

(3) the maximum absolute row sum is less than 1, therefore so is $\|\boldsymbol{G}\|_\infty$

(4) any norm less than 1 leads to convergence

$$\begin{aligned}
\boldsymbol{G} &= -\boldsymbol{D}^{-1}\boldsymbol{N} \\
&= -\begin{bmatrix}
0 & \times & \times & \times & \times & \times & \times \\
\times & \ddots & \times & \times & \times & \times & \times \\
\times & \times & \ddots & \times & \times & \times & \times \\
\frac{a_{i,1}}{a_{i,i}} & \cdots & \frac{a_{i,i-1}}{a_{i,i}} & 0 & \frac{a_{i,i+1}}{a_{i,i}} & \cdots & \frac{a_{i,m}}{a_{i,i}} \\
\times & \times & \times & \times & \ddots & \times & \times \\
\times & \times & \times & \times & \times & \ddots & \times \\
\times & \times & \times & \times & \times & \times & 0
\end{bmatrix}
\end{aligned}$$

$$\sum_{j \neq i} |a_{i,j}| < |a_{i,i}| \qquad \text{by row diagonal dominance, } \Rightarrow$$

$$\frac{1}{|a_{i,i}|}\sum_{j \neq i} |a_{i,j}| < 1$$

Since we have an arbitrary row i, this holds for any row, hence the ℓ_∞-norm is less than 1 ($\|\boldsymbol{G}\| < 1$).

Computation of Eigenvalues and Eigenvectors + Power Method

Computing the solutions of the characteristic polynomial is expensive. It takes $\Theta(n^3)$ to compute a single eigenvector of $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with Gaussian elimination. Therefore, it takes $\Theta(n^4)$ to compute all eigenvectors.

One technique used is the power method, which can be applied to find other eigenvectors. However, by itself, it converges to an eigenvector of the dominant eigenvalue. Assume the matrix $\boldsymbol{A}$ has a unique dominant eigenvalue (which has the maximum modulus), and a corresponding unique eigenvector. Take a starting vector $\boldsymbol{x^{(0)}}$, we can define the recurrence for $k = 0, 1, 2, \ldots$ as;

$$\boldsymbol{x^{(k+1)}} = \frac{\boldsymbol{A x^{(k)}}}{\|\boldsymbol{A x^{(k)}}\|}$$

In general, the idea is to keep multiplying the current vector by $\boldsymbol{A}$ and normalising. The ℓ_2-norm is useful as it is equal to the modulus of the dominant eigenvalue.

The proposition is that for a diagonalisable matrix $\boldsymbol{M} \in \mathbb{R}^{n \times n}$, with distinct eigenvalues (moduli), the iteration converges to the eigenvector of $\boldsymbol{A}$. The absolute value of the dominant eigenvalue is the limit of the normalisation constant;

$$\lim_{k \to \infty} \|\boldsymbol{A x^{(k)}}\|$$

The proof is as follows;

Let $(\boldsymbol{e_1}, \ldots, \boldsymbol{e_n})$, consisting of the unit-eigenvectors of $\boldsymbol{A}$ form a basis of $\mathbb{R}^n$. We know we can do this as $\boldsymbol{A}$ is diagonalisable. Therefore, for an arbitrary $\boldsymbol{x^{(0)}} \in \mathbb{R}^n$;

$$\boldsymbol{x^{(0)}} = \sum_{i=1}^{n} \alpha_i \boldsymbol{e_i}$$

For $k \geq 1$;

$$\begin{aligned}
\boldsymbol{x^{(k)}} &= \frac{\boldsymbol{A x^{(k-1)}}}{\|\boldsymbol{A x^{(k-1)}}\|} && \text{by recurrence relation} \\
&= \frac{\boldsymbol{A}^k \boldsymbol{x^{(0)}}}{\|\boldsymbol{A}^k \boldsymbol{x^{(0)}}\|} \\
&= \frac{\sum_{i=1}^{n} \alpha_i \lambda_i^k \boldsymbol{e_i}}{\|\boldsymbol{A}^k \boldsymbol{x^{(0)}}\|} && \text{scale by an eigenvalue of eigenvector} \\
&= \frac{\lambda_1^k}{\|\boldsymbol{A}^k \boldsymbol{x^{(0)}}\|} \sum_{i=1}^{n} \alpha_i \frac{\lambda_i}{\lambda_1}^k \boldsymbol{e_i} \\
&= \frac{\lambda_1^k}{\|\boldsymbol{A}^k \boldsymbol{x^{(0)}}\|} \left(\alpha_1 \boldsymbol{e_1} + \sum_{i=2}^{n} \alpha_i \frac{\lambda_i}{\lambda_1}^k \boldsymbol{e_i} \right) \\
\lim_{k \to \infty} \frac{\lambda_i}{\lambda_1}^k &= 0 && \text{for } i \neq 1, |\lambda_i| < |\lambda_1| \\
\lim_{k \to \infty} \boldsymbol{x^{(k)}} &= \frac{\lambda_1^k}{\|\boldsymbol{A}^k \boldsymbol{x^{(0)}}\|} \alpha_1 \boldsymbol{e_1} && \text{therefore } \boldsymbol{x^{(0)}} \to \boldsymbol{e_1}
\end{aligned}$$

Using this result, we can conclude the following;

$$\lim_{k \to \infty} \|\boldsymbol{A x^{(k)}}\| = \|\boldsymbol{A e_1}\| = \|\lambda_1 \boldsymbol{e_1}\| = |\lambda_1|$$

The limitations of this method are the following;

- starting vector $\boldsymbol{x^{(0)}}$ may have no component in the dominant eigenvector, such that $\boldsymbol{e_1}^\top \boldsymbol{x^{(0)}} = 0$, which leads to $\alpha_1 = 0$
- may have multiple eigenvalues with maximum modulus, iteration then converges to a linear combination of those eigenvectors
- convergence can be slow, if for example λ_2 is close to λ_1

Note that if the eigenvalues for the matrix $\boldsymbol{A}$ are $\lambda_1, \lambda_2, \ldots$, then the eigenvalues for the matrix $\boldsymbol{A} - \sigma \boldsymbol{I}$ are $\lambda_1 - \sigma, \lambda_2 - \sigma, \ldots$, with the same eigenvectors.

Inverse Iteration

$$\boldsymbol{Ax} = \lambda \boldsymbol{x} \Rightarrow \boldsymbol{x} = \lambda \boldsymbol{A}^{-1}\boldsymbol{x} \Rightarrow \boldsymbol{A}^{-1}\boldsymbol{x} = \tfrac{1}{\lambda}\boldsymbol{x}$$

This means that a matrix $\boldsymbol{A}^{-1}$, has eigenvalues $\frac{1}{\lambda_i}$, with the same corresponding eigenvectors as $\boldsymbol{A}$. Therefore the power iteration yields the reciprocal of the smallest eigenvalue of $\boldsymbol{A}$. An algorithm for the inverse iteration is done by a recurrence relation on pairs $(\boldsymbol{x^{(k)}}, \boldsymbol{y^{(k)}})$ for $k \geq 1$;

$$\begin{aligned} \boldsymbol{A}\boldsymbol{y^{(k+1)}} &= \boldsymbol{x^{(k)}} \\ \boldsymbol{x^{(k+1)}} &= \frac{\boldsymbol{y^{(k+1)}}}{\|\boldsymbol{y^{(k+1)}}\|} \qquad\qquad \left(= \frac{\boldsymbol{A}^{-1}\boldsymbol{x^{(k)}}}{\|\boldsymbol{A}^{-1}\boldsymbol{x^{(k)}}\|}\right) \end{aligned}$$

Since computing the inverse of $\boldsymbol{A}$ is expensive, the above can be used. We only need to solve for $\boldsymbol{A}\boldsymbol{y^{(k+1)}} = \boldsymbol{x^{(k)}}$ with any method at each iteration.

By using this with a shift σ, the eigenvalue (and eigenvector) computed will be the eigenvalue closest to σ, as that now has the smallest absolute value.

Rayleigh Quotient

Imagine we are at a point where we have an approximated eigenvector, let it be $\boldsymbol{z}$, with an approximated eigenvalue λ. Then, for $1 \leq i \leq n$, we have;

$$\sum_{j=1}^{n} a_{i,j} z_j = \lambda z_i$$

Therefore, we can see this as a least squares problem of dimension $n \times 1$;

$$\underbrace{\boldsymbol{z}}_{n\times 1}\underbrace{\lambda}_{1\times 1} = \underbrace{\boldsymbol{A}}_{n\times n}\underbrace{\boldsymbol{z}}_{n\times 1}$$

The normal equation is now $\boldsymbol{z}^\top \boldsymbol{z}\lambda = \boldsymbol{z}^\top \boldsymbol{A}\boldsymbol{z}$. Using this, we can then get the Rayleigh Quotient, which is the approximate solution;

$$\hat{\lambda} \approx \frac{\boldsymbol{z}^\top \boldsymbol{A}\boldsymbol{z}}{\boldsymbol{z}^\top \boldsymbol{z}}$$

From here, we can then choose $\sigma = \hat{\lambda}$, and shift it again. Ideally, σ will now be extremely close to the actual value, hence the shifted eigenvalue will be very small, causing the reciprocal to be very large. While σ changes each iteration, the eigenvectors don't.

Second-Dominant Eigenvalue

Deflation is a term for transforming a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ to a matrix $\boldsymbol{B} \in \mathbb{R}^{(n-1)\times(n-1)}$, with the largest (absolute) eigenvalue absent. We define a non-singular $\boldsymbol{H}$ as such, where $\boldsymbol{x_1}$ is the eigenvector of $\boldsymbol{A}$

with the largest absolute eigenvalue, α is some scalar constant, and $\boldsymbol{e_1}$ is the first standard basis vector;

$$
\begin{aligned}
\boldsymbol{H x_1} &= \alpha \boldsymbol{e_1} && \Rightarrow \\
\boldsymbol{x_1} &= \alpha \boldsymbol{H}^{-1} \boldsymbol{e_1} && \Rightarrow \\
\boldsymbol{H}^{-1} \boldsymbol{e_1} &= \frac{\boldsymbol{x_1}}{\alpha} \\
\boldsymbol{A x_1} &= \lambda_1 \boldsymbol{x_1} && \text{we know } \boldsymbol{x_1} \text{ is an eigenvector} \\
\boldsymbol{H A H}^{-1} \boldsymbol{e_1} &= \frac{1}{\alpha} \boldsymbol{H A x_1} \\
&= \frac{\lambda_1}{\alpha} \boldsymbol{H x_1} \\
&= \lambda_1 \boldsymbol{e_1} \\
&= (\boldsymbol{H A H}^{-1})_1 && \text{first column} \\
&= \begin{bmatrix} \lambda_1 & \boldsymbol{b}^\top \\ \boldsymbol{0} & \boldsymbol{B} \end{bmatrix} && \text{where } \boldsymbol{0}, \boldsymbol{b} \in \mathbb{R}^{n-1}, \boldsymbol{B} \in \mathbb{R}^{(n-1)\times(n-1)}
\end{aligned}
$$

To show they have the same eigenvalues, we do the following;

$$\boldsymbol{A x} = \lambda \boldsymbol{x} \Leftrightarrow \boldsymbol{H A x} = \lambda \boldsymbol{H x} \Leftrightarrow \boldsymbol{H A H}^{-1} (\boldsymbol{H x}) = \lambda \boldsymbol{H x}$$

Therefore, $\boldsymbol{A}$ and $\boldsymbol{H A H}^{-1}$ have the same eigenvalues, but $\boldsymbol{H A H}^{-1}$ has eigenvectors $\boldsymbol{y_i} = \boldsymbol{H x_i}$. Therefore $\boldsymbol{B}$ has eigenvectors $\lambda_2, \ldots, \lambda_n$.

We want to prove that the eigenvector corresponding to λ_2, for a given $\boldsymbol{H}$, is

$$
\begin{aligned}
\boldsymbol{x_2} &= \boldsymbol{H}^{-1} \begin{bmatrix} \beta \\ \boldsymbol{z_2} \end{bmatrix} && \text{where } \boldsymbol{z_2} \text{ is a dominant eigenvector of } \boldsymbol{B} \\
\beta &= \frac{\boldsymbol{b}^\top \boldsymbol{z_2}}{\lambda_2 - \lambda_1}
\end{aligned}
$$

To verify this, we do the following;

$$
\begin{aligned}
\boldsymbol{x_2} &= \boldsymbol{H}^{-1} \boldsymbol{y} && \boldsymbol{y} \text{ is an eigenvector of } \boldsymbol{H A H}^{-1} \text{ with eigenvalue } \lambda_2 \\
\text{let } \boldsymbol{y} &= \begin{bmatrix} y_1 \\ \boldsymbol{z} \end{bmatrix} \\
\boldsymbol{H A H}^{-1} &= \begin{bmatrix} \lambda_1 & \boldsymbol{b}^\top \\ \boldsymbol{0} & \boldsymbol{B} \end{bmatrix} \\
\begin{bmatrix} \lambda_1 & \boldsymbol{b}^\top \\ \boldsymbol{0} & \boldsymbol{B} \end{bmatrix} \begin{bmatrix} y_1 \\ \boldsymbol{z} \end{bmatrix} &= \lambda_2 \begin{bmatrix} y_1 \\ \boldsymbol{z} \end{bmatrix} && \Rightarrow \\
\lambda_2 y_1 &= \lambda_1 y_1 + \boldsymbol{b}^\top \boldsymbol{z} && \Rightarrow \\
y_1 &= \frac{\boldsymbol{b}^\top \boldsymbol{z}}{\lambda_2 - \lambda_1} \\
\lambda_2 \boldsymbol{z} &= \boldsymbol{0} y_1 + \boldsymbol{B z} \\
&= \boldsymbol{B z} && \Rightarrow \\
\beta &= y_2 \\
\boldsymbol{z_2} &= \boldsymbol{z}
\end{aligned}
$$

Householder

One of the available choices for $\boldsymbol{H}$ is the Householder transformation, which represents a reflection in the plane with a normal vector of $\boldsymbol{u}$.

$$\boldsymbol{H} = \boldsymbol{I} - \frac{2 \boldsymbol{u u}^\top}{\boldsymbol{u}^\top \boldsymbol{u}}$$

For this part, we define $\boldsymbol{u} = \boldsymbol{v} \pm \|\boldsymbol{v}\|_2 \boldsymbol{e_1}$, for a vector $\boldsymbol{v} = \boldsymbol{x_1}$. Then we have $\boldsymbol{Hv} = \mp\|\boldsymbol{v}\|\boldsymbol{e_1}$, so we find $\alpha = \mp\|\boldsymbol{x_1}\|$. This result is proved as follows;

$$
\begin{aligned}
v_1 &= \boldsymbol{v}^\top \boldsymbol{e_1} \\
\boldsymbol{u}^\top \boldsymbol{u} &= 2\|\boldsymbol{v}\|(\|\boldsymbol{v}\| \pm v_1) \\
\boldsymbol{u}^\top \boldsymbol{v} &= \boldsymbol{v}^\top \boldsymbol{u} \\
&= \|\boldsymbol{v}\|(\|\boldsymbol{v}\| \pm v_1) \\
\boldsymbol{H} &= \boldsymbol{I} - \frac{2\boldsymbol{u}\boldsymbol{u}^\top}{\boldsymbol{u}^\top \boldsymbol{u}} \qquad \Rightarrow \\
\boldsymbol{Hv} &= \boldsymbol{v} - \frac{2\boldsymbol{u}(\boldsymbol{u}^\top \boldsymbol{v})}{\boldsymbol{u}^\top \boldsymbol{u}} \\
&= \boldsymbol{v} - \boldsymbol{u} \\
&= \mp\|\boldsymbol{v}\|_2 \boldsymbol{e_1} \qquad \blacksquare
\end{aligned}
$$

26th February 2020

QR Algorithm

Note that for an upper triangular matrix

$$
\boldsymbol{U} = \begin{bmatrix}
u_{1,1} & \times & \times & \times \\
 & u_{2,2} & \times & \times \\
 & & \ddots & \times \\
 & & & u_{n,n}
\end{bmatrix}
$$

The characteristic polynomial is

$$
|\boldsymbol{U} - \lambda \boldsymbol{I}| = \prod_{i=1}^{n} (u_{i,i} - \lambda)
$$

Meaning that the eigenvalues are $u_{1,1}, u_{2,2}, \ldots, u_{n,n}$, which is the diagonal of the matrix.

Also note that a matrix $\boldsymbol{A}$ is similar to $\boldsymbol{B}$ if $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{AP}$. This is an equivalence relation, and it also preserves the characteristic polynomial, hence has the same eigenvalues.

The general idea of the QR algorithm is that it generates a sequence of similar matrices $\boldsymbol{A^{(0)}}, \boldsymbol{A^{(1)}}, \ldots$ which approaches an upper triangular matrix. The diagonal of this upper triangular matrix will then give the eigenvalues of $\boldsymbol{A}$. It follows this general sequence of steps;

(1) starting with the standard basis, set $\boldsymbol{A^{(0)}} = \boldsymbol{A}(= \boldsymbol{QR} = \boldsymbol{Q^{(1)}R^{(1)}})$

(2) apply QR decomposition with any method to compute $\boldsymbol{A^{(m)}} = \boldsymbol{Q^{(m+1)}R^{(m+1)}}$, where $\boldsymbol{Q^{(m+1)}}$ is orthogonal, and $\boldsymbol{R^{(m+1)}}$ is upper triangular

(3) define $\boldsymbol{A^{(m+1)}} = \boldsymbol{R^{(m+1)}Q^{(m+1)}} = \boldsymbol{Q^{(m+1)}}^\top \boldsymbol{A^{(m)}Q^{(m+1)}}$

note that because $\boldsymbol{Q^{(m+1)}}$ is orthogonal (actually orthonormal), $\boldsymbol{Q^{(m+1)}}^\top = \boldsymbol{Q^{(m+1)}}^{-1}$, $\boldsymbol{A^{(m+1)}}$ must be similar to $\boldsymbol{A^{(m)}}$

(4) halt after sufficient iterations

We claim that $\boldsymbol{A^{(m)}} \to \boldsymbol{U}$, where $\boldsymbol{U}$ is an upper triangular matrix, as $m \to \infty$.

$$
\boldsymbol{A^{(m)}} = \boldsymbol{Q^{(m)}}^\top \boldsymbol{Q^{(m-1)}}^\top \cdots \boldsymbol{Q^{(1)}}^\top \boldsymbol{A^{(0)}} \underbrace{\boldsymbol{Q^{(1)}Q^{(2)}} \cdots \boldsymbol{Q^{(m)}}}_{\boldsymbol{\tilde{Q}^{(m)}}} = \boldsymbol{\tilde{Q}^{(m)}}^\top \boldsymbol{A\tilde{Q}^{(m)}}
$$

Note that $\boldsymbol{\tilde{Q}^{(m)}}$ is orthogonal, as it is the product of orthogonal matrices, hence $\boldsymbol{A^{(m)}}$ must be similar to $\boldsymbol{A}$, therefore $\boldsymbol{A^{(m)}}$ has the same eigenvalues as $\boldsymbol{A}$. The eigenvectors of $\boldsymbol{A}$ are;

$$
\left\{ \boldsymbol{\tilde{Q}^{(m)}}^\top \boldsymbol{w} : \boldsymbol{w} \text{ is an eigenvector of } \boldsymbol{A^{(m)}} \right\}
$$

It is easy to compute the eigenvectors of a upper triangular matrix, which we approach, with backwards substitution. Note that it can be quicker to use the inverted shift with an approximation of the eigenvalue.

Note that for any symmetric matrix $\boldsymbol{A}$, $\boldsymbol{A^{(m)}}$ must also be symmetric. Therefore, if $\boldsymbol{A^{(m)}}$ approaches an upper triangular matrix, it must approach $\boldsymbol{\Lambda}$, which is a diagonal vector of the eigenvalues of $\boldsymbol{A}$;

$$\boldsymbol{\Lambda} = \begin{bmatrix} u_{1,1} & 0 & \cdots & 0 \\ 0 & u_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & u_{n,n} \end{bmatrix}$$

Due to this, $\boldsymbol{\tilde{Q}^{(m)}}$ approaches the matrix of eigenvectors of $\boldsymbol{A}$.

LU Decomposition

We can write any non-singular matrix $\boldsymbol{M} = \boldsymbol{LU}$. This is because we can reduce $\boldsymbol{M}$ to row echelon form (which is upper triangular), with the products of lower triangular matrices. Also note that a unit triangular matrix has all 1s on the diagonal. In this proof, we can also use the fact

$$\boldsymbol{A}^n = \boldsymbol{Q^{(1)}} \cdots \boldsymbol{Q^{(n)}} \boldsymbol{R^{(n)}} \cdots \boldsymbol{R^{(1)}}$$

We propose that $\boldsymbol{A^{(m)}} \to \boldsymbol{\Lambda}$, as $m \to \infty$, for a symmetric matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, with distinct eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0$. Also let $\boldsymbol{A} = \boldsymbol{Q\Lambda Q}^\top$ be the eigenvalue decomposition (same as spectral decomposition), and $\boldsymbol{Q}^\top = \boldsymbol{LU}$ with a **unit** lower triangular $\boldsymbol{L}$, and an upper triangular $\boldsymbol{U}$. The proof is as follows;

$$\begin{aligned}
\boldsymbol{A}^n &= \boldsymbol{Q\Lambda Q}^\top \\
&= \boldsymbol{Q^{(1)}} \cdots \boldsymbol{Q^{(n)}} \boldsymbol{R^{(n)}} \cdots \boldsymbol{R^{(1)}} && \Rightarrow \\
\boldsymbol{Q\Lambda LU} &= \boldsymbol{Q^{(1)}} \cdots \boldsymbol{Q^{(n)}} \boldsymbol{R^{(n)}} \cdots \boldsymbol{R^{(1)}} && \Rightarrow \\
\boldsymbol{Q\Lambda L} &= \boldsymbol{Q^{(1)}} \cdots \boldsymbol{Q^{(n)}} \boldsymbol{R^{(n)}} \cdots \boldsymbol{R^{(1)}} \boldsymbol{U}^{-1} && \Rightarrow \\
\boldsymbol{Q\Lambda L\Lambda}^{-n} &= \boldsymbol{Q^{(1)}} \cdots \boldsymbol{Q^{(n)}} \boldsymbol{R^{(n)}} \cdots \boldsymbol{R^{(1)}} \boldsymbol{U}^{-1} \boldsymbol{\Lambda}^{-n} \\
(\boldsymbol{\Lambda L\Lambda}^{-n})_{i,j} &= \begin{cases} \ell_{i,j} \frac{\lambda_i}{\lambda_j}^n & i > j \\ 1 & i = j \\ 0 & i < j \end{cases} && \Rightarrow \\
\lim_{n \to \infty} \boldsymbol{\Lambda L\Lambda}^{-n} &= \boldsymbol{I}
\end{aligned}$$

Therefore, $\boldsymbol{Q^{(1)}} \cdots \boldsymbol{Q^{(n)}} \boldsymbol{R^{(n)}} \cdots \boldsymbol{R^{(1)}} \boldsymbol{U}^{-1} \boldsymbol{\Lambda}^{-n} \to \boldsymbol{Q}$. Because the QR decomposition is unique, $\boldsymbol{\tilde{Q}^{(n)}} = \boldsymbol{Q^{(1)}} \cdots \boldsymbol{Q^{(n)}} \to \boldsymbol{Q}$, and $\boldsymbol{R^{(n)}} \cdots \boldsymbol{R^{(1)}} \boldsymbol{U}^{-1} \boldsymbol{\Lambda}^{-n} \to \boldsymbol{I}$, meaning $\boldsymbol{R^{(n)}} \cdots \boldsymbol{R^{(1)}} \to \boldsymbol{\Lambda}^n \boldsymbol{U}$. Note that for $\boldsymbol{\tilde{Q}^{(n)}}$ to tend to a constant matrix, $\boldsymbol{Q^{(n)}} \to \boldsymbol{I}$. From the second limit, we can argue that $\boldsymbol{R^{(n)}} \to \boldsymbol{\Lambda}$, as $\boldsymbol{\Lambda}$ is applied n times. Finally, we defined $\boldsymbol{A^{(m)}} = \boldsymbol{R^{(m)}} \boldsymbol{Q^{(m)}}$, so

$$\lim_{m \to \infty} \boldsymbol{A^{(m)}} = \boldsymbol{\Lambda}$$

. We also have $\boldsymbol{\tilde{Q}^{(m)}} \to \boldsymbol{Q}$ which diagonalises $\boldsymbol{A}$, and therefore must be the matrix of eigenvectors.

27th February 2020

Laplace Transform

The Laplace transform is a special case of the Fourier transform. The Laplace transform of a function $f(t)$, defined for real $t \geq 0$, is denoted $F(s) = (\mathcal{L}f)$;

$$F(s) = (\mathcal{L}f)(s) = \int_0^\infty e^{-st} f(t) \, \mathrm{d}t$$

Generally, $s \in \mathbb{C} = \sigma + i\omega$, therefore the Laplace transform has the "type signature" $\mathcal{L} : (\mathbb{R}^+ \to \mathbb{R}^+) \to (\mathbb{C} \to \mathbb{C})$. However, in this course we normally use positive reals, hence we have an output of $(\mathbb{R}^+ \to \mathbb{R}^+)$. Some common examples are the following;

- unit function (strictly Heaviside function, but we only consider $t \geq 0$)

$$\begin{aligned}
H(t) &= \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases} \\
f(t) &= 1 \\
F(s) &= \int_0^\infty f(t)e^{-st}\,\mathrm{d}t \\
&= \int_0^\infty e^{-st}\,\mathrm{d}t \\
&= -\frac{1}{s}\left[e^{-st}\right]_0^\infty \\
&= -\frac{1}{s}(0-1) \\
&= \frac{1}{s}
\end{aligned}$$

- exponential function (general case of the above)

$$\begin{aligned}
f(t) &= e^{-\lambda t} \\
F(s) &= \int_0^\infty e^{-\lambda t}e^{-st}\,\mathrm{d}t \\
&= \int_0^\infty e^{-(\lambda+s)t}\,\mathrm{d}t \\
&= -\frac{1}{\lambda+s}\left[e^{-(\lambda+s)t}\right]_0^\infty \\
&= -\frac{1}{\lambda+s}(0-1) \\
&= \frac{1}{\lambda+s}
\end{aligned}$$

- identity function

$$\begin{aligned}
f(t) &= t \\
F(s) &= \int_0^\infty te^{-st}\,\mathrm{d}t \\
&= \left[\frac{e^{-st}}{-s}t\right]_0^\infty - \int_0^\infty \frac{e^{-st}}{-s}\cdot 1\,\mathrm{d}t \\
&= 0 + \frac{1}{s}\int_0^\infty e^{-st}\,\mathrm{d}t \\
&= \frac{1}{s^2}
\end{aligned}$$

- exponential function

$$\begin{aligned}
f(t) &= t^n \\
F(s) &= \int_0^\infty t^n e^{-st}\,\mathrm{d}t \\
&= \left[\frac{e^{-st}}{-s}t^n\right]_0^\infty - \int_0^\infty n\frac{e^{-st}}{-s}t^{n-1}\,\mathrm{d}t \\
&= 0 + \frac{n}{s}\int_0^\infty e^{-st}t^{n-1}\,\mathrm{d}t \\
&= \frac{n}{s}\cdot\frac{n-1}{s}\int_0^\infty e^{-st}t^{n-2}\,\mathrm{d}t \\
&\vdots \\
&= \frac{n!}{s^{n+1}}
\end{aligned}$$

- non-negative real power

$$\begin{aligned}
\Gamma(z) &= \int_0^\infty t^{z-1}e^{-t}\,\mathrm{d}t \\
\Gamma(z+1) &= z\Gamma(z) && \forall z \in \mathbb{R}^+ \\
\Gamma(1) &= 1 \\
\Gamma(n) &= (n-1)! && \forall n \geq 1 \in \mathbb{N} \\
f(t) &= t^z \\
F(s) &= \frac{\Gamma(z+1)}{s^{z+1}}
\end{aligned}$$

The Laplace transform has the following properties;

(1) $\mathcal{L}$ is a bijection (unique inverse), therefore if we recognise some F, we know that f is
(2) it is a linear operator, and therefore it is simple to compute for many useful functions
(3) simple to get transform of the derivative of a function
(4) moment generating function

- Laplace transform of a convolution is a product

We can work out the derivative of a Laplace transform as such;

$$\begin{aligned}
F(s) &= \int_0^\infty e^{-st}f(t)\,\mathrm{d}t \\
\frac{\mathrm{d}}{\mathrm{d}s}F(s) &= \int_0^\infty \left(\frac{\mathrm{d}}{\mathrm{d}s}e^{-st}\right)f(t)\,\mathrm{d}t && \text{interchange of limiting operations} \\
&= \int_0^\infty -te^{-st}f(t)\,\mathrm{d}t \\
&= -\int_0^\infty (tf(t))e^{-st}\,\mathrm{d}t \\
&= -(\mathcal{L}(tf(t)))(s) \\
&= -(\mathcal{L}(\lambda t.\ tf(t)))(s) && \text{strictly with } \lambda\text{-calculus}
\end{aligned}$$

Laplace Transform of a Derivative

$$\begin{aligned}
f'(t) &= \frac{\mathrm{d}}{\mathrm{d}t} f(t) \\
(\mathcal{L}f')(s) &= \int_0^\infty e^{-st} f'(t)\,\mathrm{d}t \\
&= \left[e^{-st} f(t)\right]_0^\infty - \int_0^\infty -se^{-st} f(t)\,\mathrm{d}t \\
&= -f(0) + s\int_0^\infty e^{-st} f(t)\,\mathrm{d}t \qquad \text{assuming sub-exponential convergence for } \lim_{t\to\infty} f(t) \\
&= -f(0) + s(\mathcal{L}f)(s) \\
&= s(\mathcal{L}f)(s) - f(0)
\end{aligned}$$

Commonly, $\dot{f}$ is used to denote f differentiated with respect to t. It's important to remember the following;

$$(\mathcal{L}f')(s) = (\mathcal{L}\dot{f})(s) = sF(s) - f(0)$$

A example for the sub-exponential convergence, where it doesn't hold;

$$\begin{aligned}
f(t) &= e^{4t} \\
F(s) &= \int_0^\infty e^{-st} e^{4t}\,\mathrm{d}t \\
&= \int_0^\infty e^{(4-s)t}\,\mathrm{d}t
\end{aligned}$$

Note that $F \to \infty$ if $s \leq 4$.

Application of Laplace Transforms to Differential Equations

Starting with a fairly simple example, which would've been encountered previously;

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}t} &= \dot{y} \\
&= \mu y \\
Y &= \mathcal{L}y
\end{aligned}$$

taking Laplace transforms of both sides

$$\begin{aligned}
sY(s) - y(0) &= \mu Y(s) && \Rightarrow \\
(s-\mu)Y(s) &= y(0) && \Rightarrow \\
Y(s) &= \frac{y(0)}{s-\mu}
\end{aligned}$$

note that this is similar to $f(t) = e^{-\lambda t}$, so by uniqueness, with $\lambda = -\mu$

$$Y(s) = y(0)e^{\mu t}$$

In the harder example, the Laplace transform allows us to solve simultaneous differential equations as simultaneous linear equations.Consider these three differential equations, with the initial values $x(0) = y(0) = z(0) = 0$;

$$\begin{aligned}
\dot{x} &= z + t \\
3\dot{y} &= 2x - 3z \\
\dot{z} &= -3y
\end{aligned}$$

taking Laplace transforms of the equations, substituting initial values

$$\begin{aligned}
sX(s) - x(0) &= Z(s) + \frac{1}{s^2} && \Rightarrow \\
sX(s) - Z(s) &= \frac{1}{s^2} \\
3(sY(s) - y(0)) &= 2X(s) - 3Z(s) && \Rightarrow \\
-\frac{2}{3}X(s) + sY(s) + Z(s) &= 0 \\
sZ(s) - z(0) &= -3Y(s) && \Rightarrow \\
3Y(s) + sZ(s) &= 0
\end{aligned}$$

writing this in matrix form and solving with inverse

$$\begin{bmatrix} s & 0 & 1 \\ -\frac{2}{3} & s & 1 \\ 0 & 3 & s \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \frac{1}{s^2} \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned}
\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} &= \frac{1}{(s-1)^2(s+2)} \begin{bmatrix} s^2-3 & -3 & s \\ \frac{2s}{3} & s^2 & -\frac{3s-2}{3} \\ -2 & -3s & s^2 \end{bmatrix} \begin{bmatrix} \frac{1}{s^2} \\ 0 \\ 0 \end{bmatrix} \\
&= \frac{1}{(s-1)^2(s+2)} \begin{bmatrix} \frac{s^2-3}{s^2} \\ \frac{2}{3s} \\ -\frac{2}{s^2} \end{bmatrix} \\
&= \begin{bmatrix} -\frac{\frac{2}{3}}{(s-1)^2} + \frac{\frac{20}{9}}{s-1} + \frac{\frac{1}{36}}{s+2} - \frac{\frac{3}{2}}{s^2} - \frac{\frac{9}{4}}{s} \\ \frac{\frac{2}{9}}{(s-1)^2} - \frac{\frac{8}{27}}{s-1} - \frac{\frac{1}{27}}{s+2} + \frac{\frac{1}{3}}{s} \\ -\frac{\frac{2}{3}}{(s-1)^2} + \frac{\frac{14}{9}}{s-1} - \frac{\frac{1}{18}}{s+2} - \frac{1}{s^2} - \frac{\frac{3}{2}}{s} \end{bmatrix} \\
x(t) &= -\frac{2te^t}{3} + \frac{20e^2}{9} + \frac{e^{-2t}}{36} - \frac{3t}{2} - \frac{9}{4} \\
y(t) &= \frac{2te^t}{9} - \frac{8e^2}{27} + \frac{e^{-2t}}{27} - \frac{1}{3} \\
z(t) &= -\frac{2te^t}{3} + \frac{14e^2}{9} + \frac{e^{-2t}}{18} - t - \frac{3}{2}
\end{aligned}$$

Convolutions

The convolution of functions $f, g : \mathbb{R}^+ \to \mathbb{R}$ is denoted $f * g$, defined as;

$$\begin{aligned}
(f * g)(t) &= \int_0^t f(u)g(t-u)\,\mathrm{d}u \\
&= \int_0^t f(t-u)g(u)\,\mathrm{d}u \\
L(s) &= \mathcal{L}(f * g)(s) \\
&= \int_{t=0}^{t=\infty} e^{-st} \left(\int_{u=0}^{u=\infty} f(u)g(t-u)\,\mathrm{d}u \right) \mathrm{d}t
\end{aligned}$$

we know $t - u > 0$, therefore $u < t$, where $u, t \in \mathbb{R}^+$, now change order of integration

$$= \int_{u=0}^{u=\infty} \int_{t=u}^{t=\infty} e^{-st} f(u)g(t-u)\,\mathrm{d}u\,\mathrm{d}t$$

change t to $v = t - u$

$$\begin{aligned}
&= \int_{u=0}^{u=\infty} \int_{v=0}^{v=\infty} e^{-s(u+v)} f(u) g(v) \,\mathrm{d}u \,\mathrm{d}v \\
&= \int_{u=0}^{u=\infty} e^{-su} f(u) \,\mathrm{d}u \int_{v=0}^{v=\infty} e^{-sv} g(v) \,\mathrm{d}v \\
&= F(s)G(s)
\end{aligned}$$

Inversion

Since we know that the Laplace transform has a unique inverse, we can often solve by inspection with the use of linearity. It can be solved with an inversion formula such as the Bromwich integral (not used in our course);

$$f(t) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma - i\infty} e^{st} F(s) \,\mathrm{d}s$$

Where γ is a real greater than all the real parts of the singularities of $F(s)$ (where $F(s)$ goes to infinity). For example, a singularity of F is $-\lambda$ when

$$F(s) = \frac{1}{s + \lambda}$$

This inversion formula comes from the Laplace transform being a special case of the Fourier transform.

Dominant and Co-dominant Spaces

Suppose we have diagonalisable $\boldsymbol{M} \in \mathbb{C}^{n \times n}$, with distinct eigenvalues of moduli $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$, with corresponding eigenvectors $\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}$. Let the subspaces T_k (dominant) and U_k (co-dominant) be spanned by the bases $(\boldsymbol{v_1}, \ldots, \boldsymbol{v_k})$ (first k dominant eigenvectors) and $(\boldsymbol{v_{k+1}}, \ldots, \boldsymbol{v_n})$. They are both invariant under $\boldsymbol{A}$.

Let S be a k-dimensional subspace of $\mathbb{C}^n$ such that $S \cap U_k = \{\boldsymbol{0}\}$. Then $\boldsymbol{A}^m S \to T_k$ as $m \to \infty$. In the case $k = 1$, we have the same result as the power-iteration.

To prove this generally, we take a $\boldsymbol{v} \in S$, where $\boldsymbol{v} \neq \boldsymbol{0}$;

$$\boldsymbol{v} = \sum_{i=1}^{n} c_i \boldsymbol{v_i} = \underbrace{\sum_{i=1}^{k} c_i \boldsymbol{v_i}}_{\in T_k} + \underbrace{\sum_{i=k+1}^{n} c_i \boldsymbol{v_i}}_{\in U_k}$$

Since we know $\boldsymbol{v} \neq \boldsymbol{0}$, there has to be at least one $c_i \neq 0$, for $1 \leq i \leq k$, otherwise it would be in U_k. Applying $\boldsymbol{A}$ m times, we see the growth due to the eigenvalues as;

$$\boldsymbol{A}^m \boldsymbol{v} = \sum_{i=1}^{k} c_i \lambda_i^m \boldsymbol{v_i} + \sum_{i=k+1}^{n} c_i \lambda_i^m \boldsymbol{v_i} \Rightarrow \frac{\boldsymbol{A}^m \boldsymbol{v}}{\lambda_k^m} = \sum_{i=1}^{k} c_i \frac{\lambda_i}{\lambda_k}^m \boldsymbol{v_i} + \sum_{i=k+1}^{n} c_i \frac{\lambda_i}{\lambda_k}^m \boldsymbol{v_i}$$

However, since $\lambda_i < \lambda_k$ if $i > k$, the component in U_k approaches $\boldsymbol{0}$, when $k \to \infty$. As proven, there exists at least one non-zero c_i in the T_k component, hence $\boldsymbol{A}^m \boldsymbol{v}$ approaches a vector in T_k.

Let $\boldsymbol{Q} = [\boldsymbol{Q_1}, \boldsymbol{Q_2}]$, where $\boldsymbol{Q} \in \mathbb{C}^{n \times n}$ is formed by appending $\boldsymbol{Q_1} \in \mathbb{C}^{n \times k}$ and $\boldsymbol{Q_2} \in \mathbb{C}^{n \times (n-k)}$. Suppose the columns of $\boldsymbol{Q_1}$ form an orthonormal basis for T_k, and similar for $\boldsymbol{Q_2}$ for U_k, therefore all the columns form an orthonormal basis for $\mathbb{C}^n$. We propose the following;

$$\begin{aligned}
\boldsymbol{Q}^\top \boldsymbol{A} \boldsymbol{Q} &= \begin{bmatrix} \boldsymbol{A_{1,1}} & \boldsymbol{A_{1,2}} \\ \boldsymbol{0} & \boldsymbol{A_{2,2}} \end{bmatrix} \\
\boldsymbol{A_{1,1}} &= \boldsymbol{Q_1}^\top \boldsymbol{A} \boldsymbol{Q_1} && \in \mathbb{C}^{k \times k} \\
\boldsymbol{A_{1,2}} &= \boldsymbol{Q_1}^\top \boldsymbol{A} \boldsymbol{Q_2} && \in \mathbb{C}^{k \times (n-k)} \\
\boldsymbol{A_{2,2}} &= \boldsymbol{Q_1}^\top \boldsymbol{A} \boldsymbol{Q_2} && \in \mathbb{C}^{(n-k) \times (n-k)}
\end{aligned}$$

Subspace Convergence

Let $\boldsymbol{P(m)} = [\boldsymbol{P_1}, \boldsymbol{P_2}]$, $\boldsymbol{P(m)} \in \mathbb{C}^{n\times n}$ is formed by appending $\boldsymbol{P_1} \in \mathbb{C}^{n\times k}$ and $\boldsymbol{P_2} \in \mathbb{C}^{n\times(n-k)}$. $\boldsymbol{P(m)}$ is a unitary matrix, which is an orthogonal matrix, but allows for complex numbers. The range of $\boldsymbol{P_1}$ is $\boldsymbol{A}^m S_k$ We propose that $\boldsymbol{B_{2,1}} \to \boldsymbol{0}$, as $m \to \infty$, in the following;

$$\boldsymbol{P(m)}^\top \boldsymbol{A} \boldsymbol{P(m)} = \begin{bmatrix} \boldsymbol{B_{1,1}} & \boldsymbol{B_{1,2}} \\ \boldsymbol{B_{2,1}} & \boldsymbol{B_{2,2}} \end{bmatrix}$$

Subspace Iteration

Take a subspace S, which has no null vectors of $\boldsymbol{A}$. Suppose $(\boldsymbol{q_1^{(0)}}, \dots, \boldsymbol{q_k^{(0)}})$ is an orthonormal basis of S. Then clearly $(\boldsymbol{A}^m \boldsymbol{q_1^{(0)}}, \dots, \boldsymbol{A}^m \boldsymbol{q_k^{(0)}})$ is a basis of $\boldsymbol{A}^m S$. However, with results proven earlier, all vectors with $\boldsymbol{v_1}$-components (the dominant eigenvector of $\boldsymbol{A}$) will tend to the dominant subspace $\langle \boldsymbol{v_1} \rangle$, as $m \to \infty$. These then start to point in the same direction, thus causing the matrix to become ill-conditioned. A solution to this is to orthonormalise at each iteration of the algorithm;

- given $(\boldsymbol{q_1^{(m)}}, \dots, \boldsymbol{q_k^{(m)}})$ as a basis of S, calculate $\boldsymbol{A q_1^{(m)}}, \dots, \boldsymbol{A q_k^{(m)}}$
- choose $(\boldsymbol{q_1^{(m+1)}}, \dots, \boldsymbol{q_k^{(m+1)}})$ as a basis of $\boldsymbol{A}^{m+1} S$ by orthonormalising $\boldsymbol{A q_1^{(m)}}, \dots, \boldsymbol{A q_k^{(m)}}$ left to right, with something such as Gram-Schmidt.

This iteration holds for all subspace pairs.

Triangulation Property

We aren't expected to remember this (I hope, since I didn't write most of it). This uses the proposition in the **Subspace Convergence** section above. Since this is true for all k simultaneously, the limiting form is upper triangular.

4th March 2020

Fourier Transforms

Laplace transforms are only defined for functions on the positive reals, as e^{-st} gets very large when $t \to -\infty$, for $s > 0$. We denote the Fourier Transform of f as $\tilde{f} = \mathcal{F} f$

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega x} f(x)\,\mathrm{d}x$$

Which has the inverse transform

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega x} \tilde{f}(\omega)\,\mathrm{d}\omega$$

Dirac Delta Function

Also referred to as δ-function, which is a generalised function (or distribution) that is "undefined" at 0, and 0 everywhere else. It has an integral of 1 across the entire real line, and one can consider there to be an infinite spike at the origin.

$$\int_{-\infty}^{\infty} \delta(x)\,\mathrm{d}x = 1$$

The integral of a function multiplied by δ-function is $f(0)$, as follows (not rigorous);

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\delta(x)\,\mathrm{d}x &= \int_{-\epsilon}^{\epsilon} f(x)\delta(x)\,\mathrm{d}x && \forall \epsilon > 0 \\ &\approx f(0) \int_{-\epsilon}^{\epsilon} \delta(x)\,\mathrm{d}x && \text{at small } \epsilon \\ &= f(0) \end{aligned}$$

A normalised (the integral over the real line is 1) Gaussian function centred on the origin, with a positive constant shape parameter a takes the form

$$f(x) = \frac{1}{a\sqrt{\pi}} e^{-\frac{x^2}{a^2}}$$

Note that in the distribution $N(0, \sigma^2)$, we have $a = \sqrt{2\sigma}$. As the variance goes down towards 0, we get a "sharper" spike at the origin, which is one way we looked at the delta function;

$$\delta(x) = \lim_{a \to 0} \frac{1}{a\sqrt{\pi}} e^{-\frac{x^2}{a^2}}$$

We can define $a = \frac{1}{\sqrt{n}}$, and have the following as the limit of the delta-sequence $(\delta_n(x))$;

$$\delta(x) = \lim_{n \to \infty} \delta_n(x) = \lim_{n \to \infty} \sqrt{\frac{n}{\pi}} e^{-nx^2}$$

We can rigorously prove the following;

$$\int_{-\infty}^{\infty} \delta(x-a) f(x) \, \mathrm{d}x = \lim_{n \to \infty} \int_{-\infty}^{\infty} \delta_n(x-a) f(x) \, \mathrm{d}x = f(a)$$

Fourier Inversion Theorem

If f and $\tilde{f}$ are integrable, and f is continuous

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\omega(x-y)} f(y) \, \mathrm{d}y \, \mathrm{d}\omega$$

Using the above property (of δ), it is sufficient to show

$$\delta(y-x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega(x-y)} \, \mathrm{d}\omega$$

Letting $z = y - x$, we can just prove the following claim $\forall z \in \mathbb{R}$;

$$2\pi\delta(z) = \int_{-\infty}^{\infty} e^{-i\omega z} \, \mathrm{d}\omega$$

The proof is not examinable, but is at around Panopto timestamp 46:00.

Common Fourier Transforms

f	$\tilde{f} = \mathcal{F}f$	comment
1	$2\pi\delta(\omega)$	Heaviside function
x^n	$2\pi i^n \delta^{(n)}(\omega)$	used for polynomials
e^{iax}	$2\pi\delta(\omega - a)$	complex exponential
$e^{iax} f(x)$	$\tilde{f}(\omega - a)$	shift
$\sin ax$	$-i\pi(\delta(\omega - a) - \delta(\omega + a))$	$\sin ax = \frac{e^{iax} - e^{-iax}}{2i}$
$\cos ax$	$i\pi(\delta(\omega - a) + \delta(\omega + a))$	$\cos ax = \frac{e^{iax} + e^{-iax}}{2}$

5th March 2020

Fourier Series

Polynomial approximations (such as a Taylor series) is not suitable for a period function $f(x)$, as polynomials themselves aren't periodic. Let function $f(x)$ have a period P. We can then approximate it with a mixture of sinusoids. The base frequency is ω_0, defined as;

$$\omega_0 = \frac{2\pi}{P}$$

In general, we want to approximate $f(x)$ as a weighted mixture of harmonics, summed over n, where $n = 1, 2, \ldots$;

$$\sin \frac{2\pi}{P} nt$$

$$\cos \frac{2\pi}{P} nt$$

The result $e^{i\theta} = \cos\theta + i \sin\theta$ is also used. This can be written in three equivalent forms;

$$\begin{aligned}
f(x) &\approx f^N(x) \\
f^N(x) &= \sum_{k=-N}^{N} a_k e^{i\omega_k x} && \text{or} \\
&= \sum_{k=-N}^{N} c_k \cos\omega_k x + s_k \sin\omega_k x && \text{or} \\
&= \sum_{k=-N}^{N} d_k \cos(\omega_k x + \phi_k) \\
\omega_k &= \frac{2\pi k}{P} && a_k, c_k, d_k, s_k, \phi_k \text{ are constants}
\end{aligned}$$

Using the common results for Fourier transforms, we can observe that the Fourier transform of f^N is;

$$\mathcal{F}\left(\sum_{k=-N}^{N} a_k e^{i\omega_k x}\right) = \sum_{k=-N}^{N} a_k \mathcal{F}(e^{i\omega_k x}) = \sum_{k=-N}^{N} a_k 2\pi\delta(\omega - \omega_k) = 2\pi \sum_{k=-N}^{N} a_k \delta\left(\omega - \frac{2\pi k}{P}\right)$$

This is known as the spectrum of f. We mostly assume $x_0 = 0$, without loss of generality.

Note that $\delta_{k,n}$ is 1 if $k = n$, and 0 otherwise.

$$\begin{aligned}
\int_0^P e^{i(\omega_k - \omega_n)x}\,\mathrm{d}x &= \int_0^P e^{i\left(\frac{2\pi k}{P} - \frac{2\pi n}{P}\right)x}\,\mathrm{d}x && \text{assuming } k \neq n \\
&= \int_0^P e^{\frac{2\pi}{P}(k-n)xi}\,\mathrm{d}x \\
&= \left[\frac{P}{2\pi(k-n)} e^{\frac{2\pi}{P}(k-n)xi}\right]_0^P \\
&= \frac{P}{2\pi(k-n)}\left[e^{\frac{2\pi}{P}(k-n)xi}\right]_0^P \\
&= \frac{P}{2\pi(k-n)}(e^{2\pi i(k-n)} - 1) \\
&= \frac{P}{2\pi(k-n)}(1-1) \\
&= 0 \\
\int_0^P e^{i(\omega_k - \omega_n)x}\,\mathrm{d}x &= \int_0^P 1\,\mathrm{d}x && \text{assuming } k = n \\
&= P \\
\int_0^P e^{i(\omega_k - \omega_n)x}\,\mathrm{d}x &= P\delta_{k,n} && \text{generally}
\end{aligned}$$

This lemma can now be used to obtain a specific value for a_n as follows;

$$\begin{aligned}
f^N(x) &= \sum_{k=-N}^{N} a_k e^{i\omega_k x} && \Rightarrow \\
f^N(x)e^{-i\omega_n x} &= \sum_{k=-N}^{N} a_k e^{i\omega_k x} e^{-i\omega_n x} \\
&= \sum_{k=-N}^{N} a_k e^{i(\omega_k-\omega_n)x} && \Rightarrow \\
\int_0^P f^N(x)e^{-i\omega_n x}\,\mathrm{d}x &= \sum_{k=-N}^{N} a_k \int_0^P e^{i(\omega_k-\omega_n)x}\,\mathrm{d}x \\
&= \sum_{k=-N}^{N} a_k P\delta_{k,n} \\
&= 0+\cdots+0+a_nP+0+\cdots+0 \\
&= a_nP && \Rightarrow \\
a_n &= \frac{1}{P}\int_0^P f^N(x)e^{-i\omega_n x}\,\mathrm{d}x \\
&= \frac{1}{P}\int_0^P f^N(x)e^{-i\omega_n x}\,\mathrm{d}x && \text{if } \lim_{N\to\infty} f^N(x) = f(x)
\end{aligned}$$

For an example, let us look at the sawtooth waveform; $f(x) = \alpha x$, for $0 \le x < 1$, and extended periodically throughout the entire real line.

$$\begin{aligned}
f(x) &= \alpha x \\
P &= 1 \\
a_0 &= \frac{\alpha}{2} \\
a_n &= \frac{1}{P}\int_0^P f(x)e^{-i\omega_n x}\,\mathrm{d}x && \text{if it converges to } f \\
&= \frac{1}{1}\int_0^1 \alpha x e^{-i\frac{2\pi n}{1}x}\,\mathrm{d}x \\
&= \alpha\int_0^1 x e^{-2\pi nix}\,\mathrm{d}x \\
&= \alpha\left(\left[-\frac{1}{2\pi ni}xe^{-2\pi nix}\right]_0^1 - \int_0^1 -\frac{e^{-2\pi nix}}{2\pi ni}\,\mathrm{d}x\right) \\
&= -\frac{\alpha}{2\pi ni}\left(\left[xe^{-2\pi nix}\right]_0^1 - \underbrace{\int_0^1 e^{-2\pi nix}\,\mathrm{d}x}_{=0 \text{ since } n\neq 0}\right) \\
&= -\frac{\alpha}{2\pi ni}
\end{aligned}$$

Actually applying this, we get the following results;

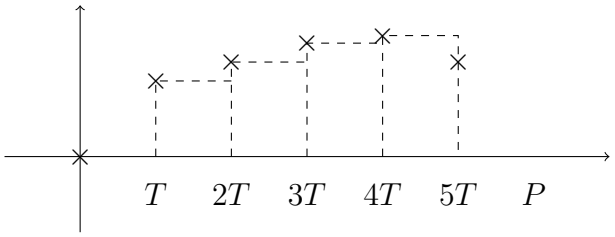
$$\begin{aligned}
f^N(x) &= \underbrace{\frac{\alpha}{2}}_{n=0} - \frac{\alpha}{2\pi i} \underbrace{\sum_{n=1}^{N} \left(\frac{1}{n} e^{i(2\pi n)x} - \frac{1}{n} e^{-i(2\pi n)x} \right)}_{n \neq 0} \\
&= \frac{\alpha}{2} - \frac{\alpha}{2\pi i} \sum_{n=1}^{N} \frac{2i}{n} \left(\frac{e^{i(2\pi n)x} - e^{-i(2\pi n)x}}{2i} \right) \\
&= \frac{\alpha}{2} - \frac{\alpha}{2\pi i} \sum_{n=1}^{N} \frac{2i}{n} \sin(2\pi n x) \\
&= \frac{\alpha}{2} - \frac{\alpha}{\pi} \sum_{n=1}^{N} \frac{1}{n} \sin(2\pi n x)
\end{aligned}$$

Sampling and Discrete Fourier Transform

Suppose the discrete function $f(x)$ is sampled at equi-spaced points $\{0, T, 2T, \ldots, (N-1)T\}$, where T divides P, and $N = \frac{P}{T}$. This is equivalent to considering the sampled function $f_s(x)$, defined as;

$$\begin{aligned}
f_s(x) &= \Psi_T(x) f(x) \\
\Psi_T(x) &= T \sum_{k=-\infty}^{\infty} \delta(x - kT)
\end{aligned}$$

This approximates rectangles of area $Tf(x)$ at points where x is a multiple of T.



The Fourier coefficients of $f_s(x)$ then become;

$$\begin{aligned}
F_k &= \frac{1}{P} \int_0^P f_s(x) e^{-i\omega_k x} \, \mathrm{d}x \\
&= \frac{1}{N} \sum_{n=0}^{N-1} f_n e^{-i\omega_k nT} \\
\omega_k &= \frac{2\pi k}{P} \\
&= \frac{2\pi k}{NT} \\
f_n &\equiv f(nT)
\end{aligned}$$

It has an inverse in the Fourier series of $f(x)$ for $x = nT$;

$$f_n = \sum_{k=0}^{N-1} F_k e^{i\omega_k nT}$$

We often assume $T = 1$, by taking some arbitrary scaling of time.

JPEG and Optics

No.

11th March 2020

Functions of Several Variables

We write functions with $n > 1$ variables as $f(x_1, \ldots, x_n)$. Let C^n be the set of functions differentiable n times. ∇ is the vector differential operator, such that

$$\nabla f(\boldsymbol{x}) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

For example, in $n = 2$, let $\boldsymbol{h} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$, then we have

$$\boldsymbol{h} \cdot \nabla f(\boldsymbol{x}) = h_1 \frac{\partial f}{\partial x_1} + h_2 \frac{\partial f}{\partial x_2}$$

This is used in the theorem for the Taylor series in n-dimensions. It states that if all all the partial derivatives of the function $f(x_1, \ldots, x_n)$ exist in a region around $x_1^{(0)}, \ldots, x_n^{(0)}$ for small enough $\boldsymbol{h}$

$$\begin{aligned}
\boldsymbol{h} &= \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \\
f(x_1^{(0)} + h_1, \ldots, x_n^{(0)} + h_n) &= f(x_1^{(0)}, \ldots, x_n^{(0)}) \\
&\quad + \boldsymbol{h} \cdot \nabla f(x_1^{(0)}, \ldots, x_n^{(0)}) \\
&\quad + \frac{1}{2!} (\boldsymbol{h} \cdot \nabla)^2 f(x_1^{(0)}, \ldots, x_n^{(0)}) \\
&\quad \vdots \\
&\quad + \frac{1}{n!} (\boldsymbol{h} \cdot \nabla)^n f(x_1^{(0)}, \ldots, x_n^{(0)}) \\
&\quad + \cdots
\end{aligned}$$

The proof is as follows;

$$\begin{aligned}
g(t) &= f(\boldsymbol{x^{(0)}} + t\boldsymbol{h}) \\
g(1) &= g(0) + g'(0) + \cdots + \frac{1}{n!} g^{(n)}(0) + \ldots
\end{aligned}$$

brief example for chain rule, if vector components are functions of t

$$\begin{aligned}
\boldsymbol{y} &= \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\
\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} \cdot \nabla &= \begin{bmatrix} \frac{\mathrm{d}y_1}{\mathrm{d}t} \\ \frac{\mathrm{d}y_2}{\mathrm{d}t} \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial}{\partial y_1} \\ \frac{\partial}{\partial y_2} \end{bmatrix} \\
&= \frac{\mathrm{d}y_1}{\mathrm{d}t} \frac{\partial}{\partial y_1} + \frac{\mathrm{d}y_2}{\mathrm{d}t} \frac{\partial}{\partial y_2}
\end{aligned}$$

using the general case, and continuing

chain rule, if components are functions of t

$$\begin{aligned}
\frac{\mathrm{d}f(\boldsymbol{y})}{\mathrm{d}t} &= \sum_{i=1}^{n} \frac{\partial f(\boldsymbol{y})}{\partial y_i} \frac{\mathrm{d}y_i}{\mathrm{d}t} \\
&= \left(\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} \cdot \nabla \right) f(\boldsymbol{y}) \\
\frac{\mathrm{d}^n f(\boldsymbol{y})}{\mathrm{d}t^n} &= \underbrace{\left(\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} \cdot \nabla \right) \cdots \left(\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} \cdot \nabla \right)}_{n \text{ times}} f(\boldsymbol{y}) \\
&= \left(\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} \cdot \nabla \right)^n f(\boldsymbol{y})
\end{aligned}$$

setting $\boldsymbol{y} = \boldsymbol{x^{(0)}} + t\boldsymbol{h}$

$$\begin{aligned}\frac{\mathrm{d}^n g}{\mathrm{d}t^n} &= \frac{\mathrm{d}^n f(\boldsymbol{y})}{\mathrm{d}t^n} \\ &= (\boldsymbol{h}\cdot\nabla)^n f(\boldsymbol{y})\end{aligned}$$

at $t = 0$, we have $\boldsymbol{y} = \boldsymbol{x^{(0)}}$

$$g^{(n)}(0) = (\boldsymbol{h}\cdot\nabla)^n(\boldsymbol{x^{(0)}})$$

$$\text{therefore } g(1) = \sum_{n=0}^{\infty} \frac{(\boldsymbol{h}\cdot\nabla)^n(\boldsymbol{x^{(0)}})}{n!}$$

Stationary Points

Similar to single variable functions, our criteria for a stationary point in $\nabla f(s_1, \ldots, s_n) = \boldsymbol{0}$, for some stationary point $\boldsymbol{s}$. Note that the slides (and I assume notes) use $\boldsymbol{x_0}$ instead of $\boldsymbol{s}$, but I feel it can be confusing. This is the same as stating

$$\frac{\partial f}{\partial x_1}(s_1, \ldots, s_n) = \cdots = \frac{\partial f}{\partial x_n}(s_1, \ldots, s_n) = 0$$

In the case where we have one variable, hence $n = 1$, there are 3 types of stationary point - maximum, minimum, and inflexion. However, in the multi-variable case, for $n > 1$, each variable can be in any of the three, therefore there are 9 possible combinations in two dimensions. This becomes exponentially more complex in more dimensions, which is why multi-variable optimisation is difficult.

In order to classify stationary points, we can use the **Hessian matrix**, defined as such;

$$\partial_i = \frac{\partial}{\partial x_i}$$

$$\begin{aligned}f_{i,j} &= \partial_i \partial_j \\ &= \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f\end{aligned} \qquad f_{i,j} = f_{j,i} \text{ for } f \in C^2$$

$$\boldsymbol{H} = \begin{bmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,n} \\ f_{2,1} & f_{2,2} & \cdots & f_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n,1} & f_{b,2} & \cdots & f_{n,n} \end{bmatrix}$$

If $\boldsymbol{H}$ is positive definite, meaning it has all positive eigenvalues and a non-zero determinant, it gives a maximum point. Similarly, if $\boldsymbol{H}$ is negative definite, it gives a maximum.

12th March 2020

Quadratic Form

The quadratic form is as follows;

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^\top \boldsymbol{x} + c$$

We can obtain the derivative of this as follows;

$$\begin{aligned}\frac{\partial x_i}{\partial x_j} &= \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \\ &= \delta_{i,j}\end{aligned}$$

$$\sum_{n}^{i=1} x_i \delta_{i,j} = x_j$$

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)$$

take arbitrary j

$$\begin{aligned}
\frac{\partial f}{\partial x_j} &= \frac{\partial}{\partial x_j}\sum_{i,k}\frac{1}{2}x_i\boldsymbol{A}_{i,k}x_k - \frac{\partial}{\partial x_j}\sum_i b_i x_i + 0 \\
&= \sum_{i,k}\frac{1}{2}\boldsymbol{A}_{i,k}\left(\frac{\partial x_i}{\partial x_j}x_k + \frac{\partial x_k}{\partial x_j}x_i\right) - \sum_i \frac{\partial x_i}{\partial x_j}b_i \\
&= \frac{1}{2}\left(\sum_k \boldsymbol{A}_{j,k}x_k + \sum_i x_i\boldsymbol{A}_{i,j}\right) - b_j \\
&= \frac{1}{2}\left(\sum_k \boldsymbol{A}_{j,k}x_k + \sum_i \boldsymbol{A}^\top_{j,i}x_i\right) - b_j \\
&= \frac{1}{2}(\boldsymbol{A}\boldsymbol{x})_j + (\boldsymbol{A}^\top\boldsymbol{x})_j - \boldsymbol{b}_j \qquad \Rightarrow \\
\nabla f &= \frac{1}{2}(\boldsymbol{A}+\boldsymbol{A}^\top)\boldsymbol{x} - \boldsymbol{b}
\end{aligned}$$

If $\boldsymbol{A}$ is symmetric, you have $\boldsymbol{A} = \boldsymbol{A}^\top$, so $\nabla f = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}$. If it is also positive definite, then the solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ globally minimises $f(\boldsymbol{x})$. Similarly, if $f(\boldsymbol{x})$ is minimised by $\boldsymbol{x} = \boldsymbol{y}$, then $\boldsymbol{y}$ is the solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, therefore there exists a duality between these two.

Iterative Solutions for Function Minimisation

We start by picking some initial vector $\boldsymbol{x_0}$ as a guess. At each step $k = 1, 2, \ldots$, we choose a direction vector $\boldsymbol{d_k}$ and a step size α_k. The iteration is

$$\boldsymbol{x_{k+1}} - \boldsymbol{x_k} + \alpha_k \boldsymbol{d_k}$$

Some steepest descent method (SDM) is as follows;

This method chooses the direction in which $f(\boldsymbol{x_k})$ decreases the fastest, hence we choose $\boldsymbol{d_k} = -\nabla f(\boldsymbol{x_k})$. When $\boldsymbol{A}$ is symmetric, $\boldsymbol{d_k} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x_k}$, in the quadratic form example, we have $\boldsymbol{d_k}$ as the residual vector.

To choose the step length, we want to minimise $f(\boldsymbol{x_k} + \alpha_k \boldsymbol{d_k})$;

$$\begin{aligned}
&\text{let } \boldsymbol{x} = \boldsymbol{x_k} \\
&\frac{\mathrm{d}f}{\mathrm{d}\alpha_k} = \sum_i \frac{\partial f(\boldsymbol{x})}{\partial x_i}\frac{\mathrm{d}x_i}{\mathrm{d}\alpha} \\
&\qquad = \sum_i \left(\nabla f(\boldsymbol{x_{k+1}})\right)_i (\boldsymbol{d_k})_i \\
&\qquad = \nabla f(\boldsymbol{x_{k+1}}) \cdot \boldsymbol{d_k} \\
&\qquad = \boldsymbol{d_{k+1}} \cdot \boldsymbol{d_k}
\end{aligned}$$

To minimise this, set it to zero, therefore we want to choose α such that $\boldsymbol{d_{k+1}}$ is orthogonal to $\boldsymbol{d_k}$. Therefore in general we want to solve $(\nabla f(\boldsymbol{x_k})) \cdot (\nabla f(\boldsymbol{x_{k+1}})) = 0$

An example is as follows (quadratic form);

$$\begin{aligned}
\boldsymbol{r_k}^\top(\boldsymbol{b} - \boldsymbol{A}(\boldsymbol{x_k} + \alpha_k \boldsymbol{r_k})) &= \boldsymbol{r_k}^\top(\boldsymbol{r_k} - \alpha_k \boldsymbol{A}\boldsymbol{r_k}) \\
&= 0 \qquad \text{(solve)} \\
\alpha_k &= \frac{\boldsymbol{r_k}^\top \boldsymbol{r_k}}{\boldsymbol{r_k}^\top \boldsymbol{A}\boldsymbol{r_k}} \\
&= \frac{\|\boldsymbol{r_k}\|^2}{\boldsymbol{r_k}^\top \boldsymbol{A}\boldsymbol{r_k}}
\end{aligned}$$

By the following iteration, we approximate a solution for $\boldsymbol{x}$;

$$\begin{aligned}
\boldsymbol{x_k} &= \boldsymbol{A r_k} \\
\alpha_k &= \frac{\|\boldsymbol{r_k}\|^2}{\boldsymbol{r_k}^\top \boldsymbol{v_k}} \\
\boldsymbol{x_{k+1}} &= \boldsymbol{x_k} + \alpha_k \boldsymbol{r_k} \\
\boldsymbol{r_{k+1}} &= \boldsymbol{r_k} - \alpha_k \boldsymbol{v_k}
\end{aligned}$$

This takes a single matrix-vector multiplication per iteration, and one more in the initialisation.

Example of Iterative Method

Suppose we want to minimise the function $f(x, y) = x^2 + y^2$. By inspection, we know the minimum, for real x, y, is $(0, 0)$. However, numerically, we do the following;

$$\begin{aligned}
\text{guess } \boldsymbol{x_0} &= (x_0, y_0) && \text{minimum} \\
\boldsymbol{d_0} &= -\nabla f(x_0, y_0) \\
&= -(2x, 2y)|_{\boldsymbol{x}=\boldsymbol{x_0}} \\
&= -2(x_0, y_0) \\
&= -2\boldsymbol{x_0} \\
x_1 &= x_0 + \alpha(-2x_0) \\
&= (1 - 2\alpha)x_0 \\
y_1 &= y_0 + \alpha(-2y_0) \\
&= (1 - 2\alpha)y_0 \\
\boldsymbol{d_1} &= -2(x_1, y_1) \\
&= -2\boldsymbol{x_1} \\
\text{solve } \boldsymbol{d_0} \cdot \boldsymbol{d_1} &= 0 && \Leftrightarrow \\
x_0 x_1 + y_0 y_1 &= 0 && \Leftrightarrow \\
(1 - 2\alpha)x_0^2 + (1 - 2\alpha)y_0^2 &= 0 && \Leftrightarrow \\
(1 - 2\alpha)(x_0^2 + y_0^2) &= 0 && \Rightarrow \\
\alpha &= \frac{1}{2}
\end{aligned}$$

after this computation, we have the next step, so;

$$\begin{aligned}
\boldsymbol{d_0} &= -2(x_0, y_0) \\
&= -2\boldsymbol{x_0} \\
\boldsymbol{x_1} &= \boldsymbol{x_0} + \alpha \boldsymbol{d_0} \\
&= \boldsymbol{x_0} + \frac{1}{2}(-2\boldsymbol{x_0}) \\
&= \boldsymbol{0} \\
&= (0, 0)
\end{aligned}$$

This algorithm converged in a single step, due to the symmetry (hence no "zig-zagging" needed).