Week 11 (lecture 1)

What we'll cover

[A] Spectral theorem for symmetric matrices

__Theorem__ Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, i.e. $A^T = A$. Then

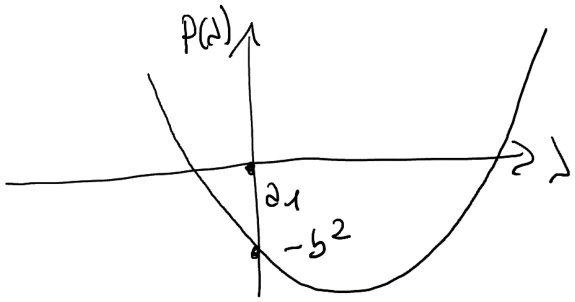
- there exists an orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ (i.e. $Q^T Q = I$) such that $Q^T A Q = D$ where D is a diagonal matrix. Thus it is possible to find a set of orthonormal eigenvectors (the columns of Q).
- All entries of D (i.e. the eigenvalues of A) are real
- Eigenvectors corresponding to different eigenvalues are orthogonal

__Example__ Let us start from $n=2$

$$A = \begin{pmatrix} a_1 & b \\ b & a_2 \end{pmatrix} \quad \text{so} \quad A^T = A$$

$$p_A(\lambda) = (a_1-\lambda)(a_2-\lambda) - b^2 = \lambda^2 - (a_1+a_2)\lambda + a_1 a_2 - b^2$$

↳ Plot



If $b=0$ the matrix is already diagonal, otherwise there are two real solutions. Consider λ_1 and $\underline{x}$ ($A\underline{x} = \lambda_1 \underline{x}$) is the corresponding eigenvector.

Consider $\underline{y} \in [\text{span}(\underline{x})]^{\perp}$ i.e. a vector orthogonal to $\underline{x}$. Then $A\underline{y}$ is also orthogonal to $\underline{x}$:

$$\underline{x} \cdot A\underline{y} = (x_1, x_2) \begin{pmatrix} a_1 & b \\ b & a_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} =$$

$$\left(\begin{pmatrix} a_1 & b \\ b & a_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right)^T \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \overset{\underline{x} \text{ is an eigen.}}{=} \left(\lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right)^T \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} =$$

$$= \lambda\, \underline{x} \cdot \underline{y} = 0 \rightsquigarrow \underline{x} \text{ and } \underline{y} \text{ are orthog.}$$

Since $\left[\text{Span}(\underline{x})\right]^{\perp}$ is 1-dimensional $A\underline{y}$ must be proportional to y, i.e. $\underline{y}$ is an eigenvector which has to correspond to the other solution λ_2 of $P_A(\lambda) = 0$

We can normalise $\underline{x}, \underline{y}$ by defining

$$\underline{u}_1 = \frac{\underline{x}}{\|\underline{x}\|} \quad \text{and} \quad \underline{u}_2 = \frac{\underline{y}}{\|\underline{y}\|} \quad \text{so}$$

$$Q = \begin{pmatrix} (u_1)_1 & (u_2)_1 \\ (u_1)_2 & (u_2)_2 \end{pmatrix} \quad \text{is orthogonal and}$$

$$Q^T A Q = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = D$$

Proof: A key point to generalise the argument above to the $n \times n$ case is to show that any symmetric matrix has at least 1 real eigenvalue and so an eigenvector in $\mathbb{R}^n$. Suppose that this can be done (see below) and the corresponding eigenvector is $\underline{v}_1$, i.e. $A\underline{v}_1 = \lambda_1 \underline{v}_1$. The interesting case is when $\underline{v}_1 \neq \alpha_i \underline{x}_i$ for any $i = 1, \dots n$ with $\alpha_i \in \mathbb{R}$ and

$$\underline{X} = \left\{ \underline{x}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \underline{x}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right\}$$

If $\underline{v}_1 = \alpha_i \underline{x}_i$ for some i the matrix has a diagonal block, for instance if $\underline{v}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, we have

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \\ 0 & a_{2n} & \dots & a_{nn} \end{pmatrix}$$

The first column must be prop. to $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ to satisfy the eigenvector/eigenvalue equation with the above $\underline{v}_1$ and the first row is fixed by $A^T = A$. So we can focus on the $(n-1) \times (n-1)$ block which defines a linear map from $(\mathrm{span}(\underline{v}_1))^\perp$ to itself.

Then we can change basis to an ortonormal basis B containing $\underline{u}_1 = \dfrac{\underline{v}_1}{\|\underline{v}_1\|}$ (which can be constructed by using $\underline{v}_1$ and $(n-1)$ $\underline{x}_i$ in the Gram-Schmidt process). Again in the basis B_1 A is block diagonal

$$[A]^B_B = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & b_{11} & \dots & b_{1n-1} \\ \vdots & \vdots & & \\ 0 & b_{1n-1} & & b_{n-1n-1} \end{pmatrix}$$

and we can focus on the $(n-1) \times (n-1)$ subblock and repeat the same argument till we construct the basis that diagonalises A. Note that, as in the 2x2 case, the symmetry property ensures that

if $\underset{\sim}{v}_1 \cdot \underline{v} = 0$ the $\underline{v}_1 \cdot (A\underline{v}) = 0$

$\Downarrow$

$$[\underline{v}]_B = \begin{pmatrix} 0 \\ \beta_1 \\ \vdots \\ \beta_{n-1} \end{pmatrix} \quad \text{and} \quad [A]_B^B \, [\underline{v}]_B = \begin{pmatrix} 0 \\ * \\ \times \\ \vdots \\ * \end{pmatrix}$$

By construction the new eigenvector $\underset{\sim}{u}_2$ of $[A]_B^B$ obtained by looking at the sub-block $\begin{pmatrix} b_{11} & \cdots & b_{1\,n-1} \\ \vdots & & \\ b_{n-1\,1} & \cdots & b_{n-1\,n-1} \end{pmatrix}$ belongs to

$(\text{span}(\underline{u}_1))^{\perp}$ and so, after normalising it, $\underset{\sim}{u}_2 \cdot \underline{u}_1 = 0$.

If $\lambda_1 \neq \lambda_2$ we can see the orthogonality property also as follows

$$\left.\begin{array}{l} A\,\underline{u}_1 = \lambda_1 \underline{u}_1 \\ A\,\underline{u}_2 = \lambda_2 \underline{u}_2 \end{array}\right\} \quad \begin{array}{l} \underline{u}_2 \cdot (A\,\underline{u}_1) = \lambda_1\, \underline{u}_2 \cdot \underline{u}_1 \\ \quad \shortparallel \\ (A^T \underline{u}_2) \cdot \underline{u}_1 = (A\,\underline{u}_2) \cdot \underline{u}_1 = \lambda_2\, \underline{u}_2 \cdot \underline{u}_1 \end{array}$$

If $\lambda_1 \neq \lambda_2$ this is possible only if $\underline{u}_2 \cdot \underline{u}_1 = 0$

In any case the basis of eigenvectors $\underline{u}_i$ constructed above is orthonormal so Q is orthogonal.

<u>Loose end</u>: in order to prove that there alway exists a real eigenvalue one can

- generalise the problem to the complex domain and use the fundamental theorem of algebra. This ensures the existence of a (possibly complex) eigenvalue.

Suppose that indeed $\lambda = (\mathrm{Re}\,\lambda) + i(\mathrm{Im}\,\lambda)$ the

$$A\cdot(\underline{x}+i\underline{y}) = (\mathrm{Re}\,\lambda + i\,\mathrm{Im}\,\lambda)(\underline{x}+i\underline{y}) \Rightarrow$$

c.c.

$$A\cdot(\underline{x}-i\underline{y}) = (\mathrm{Re}\,\lambda - i\,\mathrm{Im}\,\lambda)(\underline{x}-i\underline{y}) \quad \text{i.e.}$$

λ^* is another eigenvalue. Then

as $\lambda \neq \lambda^*$ are different eigenvalues

$(\underline{x}-i\underline{y})\cdot(\underline{x}+i\underline{y}) = \underline{x}\cdot\underline{x} + \underline{y}\cdot\underline{y}$ should vanish but this is possible only if $\underline{x}=\underline{y}=0$. Thus a complex eigenvalue cannot exist and the λ obtained from the fundamental theorem of algebra must be real.

<u>Comment</u>: if A, B are symmetric $\mathbb{R}^{n\times n}$ matrices satisfying $AB = BA$, there exists an orthogonal matrix Q that diagonalises <u>both</u> A and B. <u>Hint</u> of a proof: start by showing that if $\underline{v} \in N(A-\lambda I)$ then $B\underline{v} \in N(A-\lambda I)$ ...

What we'll cover (lectures 2-3)

A] Example of $Q^T A Q = D$ with $A = A^T$

B] Orthogonal projection

C] Least square problem

D] Hom spaces and End algebras ← (non-ex)

A] Let $A = \begin{pmatrix} 0 & 2 & -1 \\ 2 & 3 & -2 \\ -1 & -2 & 0 \end{pmatrix}$: find Q and P.

$$P_A(\lambda) = \det \begin{pmatrix} -\lambda & 2 & -1 \\ 2 & 3-\lambda & -2 \\ -1 & -2 & -\lambda \end{pmatrix} = \lambda^2(3-\lambda) + 4\lambda + 4 + 4\lambda + 4 + (\lambda - 3)$$

$$P_A(\lambda) = 0 \Rightarrow -\lambda^3 + 3\lambda^2 + 9\lambda + 5 = 0$$

$\lambda = -1$ is an "easy" root. It has also multiplicity 2 (the discriminant is of $P_A(\lambda)$ vanishes). We have

$$P_A(\lambda) = (\lambda+1)(-\lambda^2 + a\lambda + b) \Rightarrow b = 5, \ a = 4 \text{ so}$$

$$P_A(\lambda) = (\lambda+1)\underbrace{(-\lambda^2+4\lambda+5)}_{(\lambda+1)(-\lambda+5)} = -(\lambda+1)^2(\lambda-5)$$

Thus we have two eigenvalues $\lambda_1 = -1$, $\lambda_2 = 5$

The eigenvectors of λ_1 are in the space

$$N(A-\lambda_1 I) = N\begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 1 \end{pmatrix}.$$ This ~~space~~

is two dimensional and an a basis is

$$\underline{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \underline{x}_2 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}.$$ We can make

it orthonormal à la Gram-Schmidt

$$\underline{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}; \quad \underline{v}_2 = \underline{x}_2 - (\underline{x}_2 \cdot \underline{u}_1)\,\underline{u}_1 \Rightarrow$$

$$\underline{v}_2 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} - \left((-2\ 1\ 0)\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right)\left(\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right)$$

$$\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$ Thus

$$\underline{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}; \qquad \underline{u}_2 = \frac{1}{\sqrt{3}}\begin{pmatrix}-1\\1\\1\end{pmatrix}$$

The eigenvector of λ_2 is in the space

$$N\begin{pmatrix}-5 & 2 & -1\\ 2 & -2 & -2\\ -1 & -2 & -5\end{pmatrix} \Rightarrow \left(\begin{array}{ccc|c}-5 & 2 & -1 & 0\\ 2 & -2 & -2 & 0\\ -1 & -2 & -5 & 0\end{array}\right) \overset{REF}{\sim} \left(\begin{array}{ccc|c}-5 & 2 & -1 & 0\\ 1 & -1 & -1 & 0\\ 0 & -3 & -6 & 0\end{array}\right) \sim$$

$$\left(\begin{array}{ccc|c}-5 & 2 & -1 & 0\\ 0 & -3 & -6 & 0\\ 0 & -3 & -6 & 0\end{array}\right) \sim \left(\begin{array}{ccc|c}-5 & 2 & -1 & 0\\ 0 & 1 & 2 & 0\\ 0 & 0 & 0 & 0\end{array}\right) \Rightarrow \underline{x}_3 = \begin{pmatrix}-1\\-2\\1\end{pmatrix}$$

We have $\underline{u}_3 = \dfrac{\underline{x}_3}{\|\underline{x}_3\|} = \dfrac{1}{\sqrt{6}}\begin{pmatrix}-1\\-2\\1\end{pmatrix}$. By the theorem

proved above $\underline{u}_1 \cdot \underline{u}_3 = \underline{u}_2\,\underline{u}_3 = 0$. Check

$$\underline{u}_1 \cdot \underline{u}_3 \sim (1\;0\;1)\begin{pmatrix}-1\\-2\\1\end{pmatrix} = 0 \quad ✓$$

$$\underline{u}_2\,\underline{u}_3 \sim (-1\;1\;1)\begin{pmatrix}-1\\-2\\1\end{pmatrix} = 1 - 2 + 1 = 0 \quad ✓$$

$$Q = \begin{pmatrix}1/\sqrt{2} & -1/\sqrt{3} & -1/\sqrt{6}\\ 0 & 1/\sqrt{3} & -2/\sqrt{6}\\ 1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6}\end{pmatrix} \quad \text{and} \quad Q^T A Q = \begin{pmatrix}-1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 5\end{pmatrix}$$

B] Orthogonal decomposition

<u>Theorem</u>: Let H be a vector subspace of $\mathbb{R}^n$. Then $\forall\, \underline{y} \in \mathbb{R}^n$ can be decomposed uniquely as follows

$$\underline{y} = \hat{\underline{y}} + \underline{z} \quad \text{with} \quad \begin{matrix} \hat{\underline{y}} \in H \\ \underline{z} \in H^{\perp} \end{matrix}$$

Proof: Choose an orthonormal basis $\underline{u}_i$ for H.

The $\hat{\underline{y}} = \sum_i (\underline{y} \cdot \underline{u}_i)\, \underline{u}_i \in H$ (obvious)

$\underline{z} = \underline{y} - \hat{\underline{y}}$ is such that $\underline{z} \cdot \underline{u}_j = \underline{y} \cdot \underline{u}_j - \hat{\underline{y}}\, \underline{u}_j =$

$$= (\underline{y} \cdot \underline{u}_j) - \sum_i (\underline{y} \cdot \underline{u}_i) \underbrace{(\underline{u}_i \cdot \underline{u}_j)}^{\delta_{ij}} = 0 \Rightarrow \underline{z} \in H^{\perp}$$

Uniqueness: suppose $\underline{y} = \hat{\underline{y}} + \underline{z} = \hat{\underline{y}}_1 + \underline{z}_1 \Rightarrow$

$\underline{w} \equiv \hat{\underline{y}} - \hat{\underline{y}}_1 = \underline{z} - \underline{z}_1$ and by what we just proved

it is a vector belonging to both H and

$H^\perp$. Thus $\underline{w} \cdot \underline{w} = 0 \Rightarrow \underline{w} = 0$ so

$\hat{\underline{y}} = \hat{\underline{y}}_1$, $\underline{z} = \underline{z}_1$ ✓

Thus we can define a linear operator

$\text{proj}_H(\underline{v}) = \sum (\underline{v} \cdot \underline{u}_i)\, \underline{u}_i$ (where $\{\underline{u}_i\}$ is an orthonormal basis for H as above) from $\mathbb{R}^n$ to H. proj_H is called projector and $\text{proj}_H(v)$ is the orthogonal projection of $\underline{v}$ onto H.

<u>Theorem</u>: In the same setting as above

$$\hat{\underline{y}} = \text{proj}_H(\underline{y}) \Rightarrow \|\underline{y} - \hat{\underline{y}}\| \leq \|\underline{y} - \underline{v}\| \quad \forall\, \underline{v} \in H$$

By using $\underline{y} = \hat{\underline{y}} + \underline{z}$ we have

$$\|\underline{y} - \hat{\underline{y}}\|^2 = \underline{z} \cdot \underline{z} \overset{?}{\leq} \|(\hat{\underline{y}} + \underline{v}) + \underline{z}\|^2 = (\underline{y} + \underline{v}) \cdot (\underline{y} + \underline{v}) + \underline{z} \cdot \underline{z}$$

$$\Rightarrow \|\underline{y} + \underline{v}\|^2 \geq 0 \quad \text{which is true ✓}$$

Corollorary H is a vector subspace of V. Then

$$\dim(H^{\perp}) + \dim(H) = \dim(V)$$

Course work

Ex1] Let $H = \text{span}\left\{\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}\right\}$. Find the best approximation of $\underset{\sim}{y} = \begin{pmatrix}2\\-1\\1\end{pmatrix}$ in H.

We have $H = \text{span}\left\{\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix} - \begin{pmatrix}1\\1\\0\end{pmatrix}\right\} = \left\{\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$

so the basis is orthogonal and $H = \text{span}\left\{\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$ with an orthonormal basis. Then

$$P_H(\underset{\sim}{y}) = (\underset{\sim}{y}\ \underline{u}_1)\,\underline{u}_1 + (\underline{y}\ \underline{u}_2)\,\underline{u}_2 = \frac{1}{2}\begin{pmatrix}1\\1\\0\end{pmatrix} + \begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}1/2\\1/2\\1\end{pmatrix}$$

c] Least squares problem

Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$

- The least squares problem for the system $A\underline{x} = \underline{b}$ is to find $\underline{x} \in \mathbb{R}^n$ to manes $\|\underline{b} - A\underline{x}\|$ as small as possible
- A least squares solutions of $A\underline{x} = \underline{b}$ is a $\hat{\underline{x}} \in \mathbb{R}^n$ such that $\|\underline{b} - A\hat{\underline{x}}\| \leq \|\underline{b} - A\underline{x}\|\ \forall\, \underline{x} \in \mathbb{R}^n$

Theorem Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$

The set of least squares solutions of $A\underline{x} = \underline{b}$ is formed by the solutions of

$$A^T A \underline{x} = A^T \underline{b}$$ ← normal equation

Proof: For any $\underline{x} \in \mathbb{R}^n$, $A \cdot \underline{x} \in Col(A)$ (i.e. it is part of the column space of A). Of course also $proj_{col(A)} b \in col(A)$ (we use $\hat{\underline{b}}$ to indicate $proj_{col(A)} b$). To minimise $\|\underline{b} - A\underline{x}\|$ we need to choose a $\hat{x}$ such that $\hat{\underline{b}} = A\hat{\underline{x}}$: in this way we kill the contribution to $\|\underline{b} - A\hat{\underline{x}}\|$ in $col(A)$ and we are left just with that from the component in $(col(A))^\perp$ (see part B]). Thus $b - \hat{b} \in (col(A))^\perp$

Thus a least squares solution satisfies

$$b - \hat{b} \in (col(A))^\perp \underset{\text{week 10}}{=} N(A^T) \Rightarrow A^T(b - \underset{(1)}{\hat{b}}) = 0 \Rightarrow$$

$$\underset{(1)}{A^T A \hat{x}} = \underset{(2)}{A^T b} \qquad \checkmark$$

Example: Consider the matrix

$A = \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix}$ and $\underline{b} = \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix}$. The system $A\underline{x} = \underline{b}$ does not have solutions. Let us find the best approximate solution by using the theorem above.

$$A^T A = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 17 & 1 \\ 1 & 5 \end{pmatrix}$$

$$A^T b = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix} = \begin{pmatrix} 19 \\ 11 \end{pmatrix} \Rightarrow$$

$$\left(\begin{array}{cc|c} 17 & 1 & 19 \\ 1 & 5 & 11 \end{array}\right) \sim \left(\begin{array}{cc|c} 17 & 1 & 19 \\ 0 & -84 & -168 \end{array}\right) \sim \left[\begin{array}{cc|c} 17 & 1 & 19 \\ 0 & 1 & 2 \end{array}\right] \sim$$

$$\left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array}\right) \Rightarrow \hat{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

D] Hom spaces and End algebras (non ex.)

<u>Definition</u> Let U and V be vector spaces $\mathrm{Hom}(U, V)$ is the set of all linear transformation from U to V. $\mathrm{Hom}(U, V)$ is itself a vector space.

If you choose coordinates in U and V, You can see that $Hom(U, V)$ is isomorphic to $\mathbb{R}^{m \times n}$ with $\dim(U) = \mathbb{R}^n$, $\dim(V) = \mathbb{R}^m$

If the initial and the final space are the same we call the set of all linear transformations an endomorphism $Hom(V, V) = End(V)$.

It is isomorphic to the square matrices.

If we restrict to the set of linear maps that also isomorphism then we have $Aut(V)$ which itself is isomorphic to the set of square invertible matrices.