## MATH 125 — Lecture 13

### * circle theorem

Let $A$ be an $n \times n$ matrix
$x$ is an eigenvector of $A$ with corresponding eigenvalue $\lambda$ i.e. $Ax = \lambda x$

Let $x_i$ be the largest entry of $x$ in absolute magnitude. Now consider the $i$-th entry of $Ax$

$$(Ax)_i = (\lambda x)_i$$

$$(Ax)_i = \underbrace{i\begin{pmatrix} \rule{3cm}{0.4pt} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} | \end{pmatrix}}_{x} = \sum_j A_{ij} x_j$$

$$\sum_{j=1}^{n} A_{ij} x_j = \lambda x_i$$

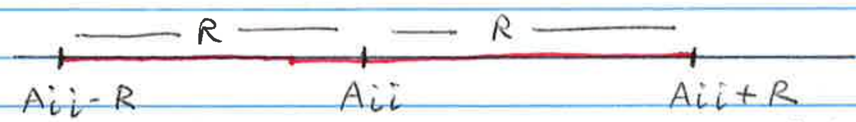
$$\sum_{j \neq i} A_{ij} x_j + A_{ii} x_i = \lambda x_i$$

$$\sum_{j \neq i} A_{ij} x_j = (\lambda - A_{ii}) x_i$$

$$|\lambda - A_{ii}| = \left| \frac{1}{x_i} \sum_{j \neq i} A_{ij} x_j \right| \leq \sum_{j \neq i} |A_{ij}| \underbrace{\left| \frac{x_j}{x_i} \right|}_{\leq 1 \text{ by definition of } x_i} \leq \sum_{j \neq i} |A_{ij}| \cancel{\frac{x_j}{x_i}} = \sum_{j \neq i} |A_{ij}|$$

Therefore, $|\lambda - A_{ii}| \leq \sum_{j \neq i} |A_{ij}|$

$$R = \sum_{j \neq i} |A_{ij}|$$

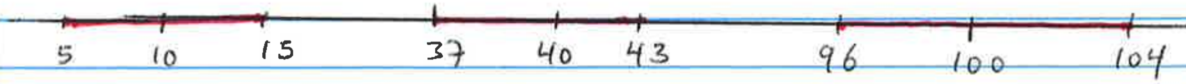



The eigenvalue $\lambda$ is in the region $[A_{ii} - R, A_{ii} + R]$

**Gerschgorin circle theorem** Every eigenvalue of $A$ lies within atleast one of the circles centered at $A_{ii}$ with radius $R_i = \sum_{j \neq i} |A_{ij}|$

**Example** Let $A = \begin{bmatrix} 100 & -1 & 3 \\ -1 & 40 & -2 \\ 3 & -2 & 10 \end{bmatrix}$

$A_{11} = 100$ ; $R_1 = |-1| + 3 = 4$
$A_{22} = 40$ ; $R_2 = |-1| + |-2| = 3$
$A_{33} = 10$ ; $R_3 = |3| + |-2| = 5$




Minimum eigenvalue $\geq 5$
Maximum eigenvalue $\leq 104$
The eigenvalues of A lie in the red regions

Note, we can apply the circle theorem to show that the matrix which solves the linear system for splines is invertible.

Proof The Greschgorin disks are shown below


The minimum eigenvalue is atleast 1.
A does not contain a zero eigenvalue
$\Rightarrow$ A is invertible

---

Least squares

- In interpolation, we require the interpolant to exactly pass through the given data.
- In practice, observed data is prone to measurement error.
- In least squares, we relax the requirement of exactly passing through data points.

Data $(t_1, y_1)$ $(t_2, y_2)$ ... $(t_n, y_n)$
Let's say we want "to fit" a line $p(t) = a + bt$ to the data
we define the following error function.

$$E = \sum_{i=1}^{n} [y_i - p(t_i)]^2 = \underbrace{\sum_{i=1}^{n} [y_i - (a+bt_i)]^2}_{\text{Least squares error}}$$

$$E = \left\| \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{\vec{y}} - \begin{pmatrix} a+bt_1 \\ a+bt_2 \\ \vdots \\ a+bt_n \end{pmatrix} \right\|_2^2 = \left\| \vec{y} - \underbrace{\begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} a \\ b \end{pmatrix}}_{\vec{x}} \right\|_2^2 = \| \vec{y} - A\vec{x} \|_2^2$$

(Here on, we drop the $\rightarrow$ for ease)

## Two cases

① $E = 0$

$\Rightarrow$ This happens only if all data points lie on the same line

$\Rightarrow$ This is less likely to happen if one collects real data

② $E \neq 0$

The goal is to make $y$ close to $Ax$ as small as possible

$$\text{Range}(A) = \{ Ax \,;\, x \in \mathbb{R}^n \} \qquad A \in \mathbb{R}^{m \times n}$$

With that, our aim is to find a vector in Range(A) that is closest to $y$

**Best approximation theorem** Let $W$ be a subspace of $\mathbb{R}^n$. Let $\hat{y}$ be orthogonal projection of $y$ onto $W$. Then $\hat{y}$ is the closest point in $W$ to $y$. Formally,

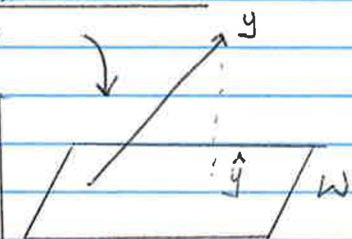
$$\| y - \hat{y} \| < \| y - z \| \quad \text{for all } z \neq \hat{y}$$

**Proof** $\quad y - z = \underline{y - \hat{y}} + \underline{\hat{y} - z}$

$$\begin{aligned} \| y - z \|^2 &= \| y - \hat{y} + \hat{y} - z \|^2 \\ &= (y - \hat{y} + \hat{y} - z)^T (y - \hat{y} + \hat{y} - z) \\ &= (y - \hat{y})^T (y - y) + (\hat{y} - z)^T (\hat{y} - z) \\ &\quad + 2 (y - \hat{y})^T (\hat{y} - z) \\ &= \| y - \hat{y} \|^2 + \| \hat{y} - z \|^2 + 2 \underbrace{(y - \hat{y})^T (\hat{y} - z)}_{= 0} \\ &> \| y - \hat{y} \|^2 \quad \text{for all } z \neq \hat{y} \end{aligned}$$

Note: $\|z\|^2 = \langle z, z \rangle = z^T z$ standard inner product

lies in $W$




**Remark** This is a very important result and holds in the general setting where $W$ is a subspace of an inner product space

Note: By definition of orthogonal projection, $y - \hat{y} \perp p \quad p \in W$

Let's apply best approximation theorem to least squares

$\Rightarrow$ To find $\hat{y}$, we need to project $y$ orthogonally onto Range$(A)$.

Note $y-\hat{y} \perp z$ $\quad z \in$ Range$(A)$

$$\therefore \quad \begin{matrix} y-\hat{y} \perp a_1 \\ y-\hat{y} \perp a_2 \\ \vdots \\ y-\hat{y} \perp a_3 \end{matrix} \quad \Longrightarrow \quad \begin{matrix} a_1^T(y-\hat{y}) = 0 \\ a_2^T(y-\hat{y}) = 0 \\ \vdots \\ a_n^T(y-\hat{y}) = 0 \end{matrix}$$

compactly, we have

$$\begin{pmatrix} — a_1^T — \\ — a_2^T — \\ \vdots \\ — a_n^T — \end{pmatrix} (y-\hat{y}) = \vec{0}$$

$A^T y = A^T \hat{y}$ $\qquad$ Note $\hat{y} \in$ Range$(A)$ i.e $\hat{y} = Ax$ for some $x$

$A^T y = A^T A x$

$\boxed{A^T A x = A^T y}$ $\qquad$ Normal equations

---

What if we want to fit $(t_1, y_1)$ $(t_2, y_2) \cdots (t_n, y_n)$ to $p(t) = a + bt + ct^2$?

$\Rightarrow$ $A = \begin{pmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ & \vdots & \\ 1 & t_n & t_n^2 \end{pmatrix}$ $\qquad x = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ $\qquad$ solve $A^T A x = A^T y$

The only thing that changes is $A$. Essentially, we still have a linear system.

<u>Exercise</u> Fit $(t_1, y_1)$ $(t_2, y_2) \cdots (t_n, y_n)$ to $y = a e^{-t} + b$

$A = \begin{pmatrix} e^{-t_1} & 1 \\ e^{-t_2} & 1 \\ \vdots & \\ e^{-t_n} & 1 \end{pmatrix}$ $\qquad x = \begin{pmatrix} a \\ b \end{pmatrix}$

Note we can generalize least squares to spaces of functions

<u>Example</u> Let $P_1 = 1$, $P_2 = x$ and $P_3 = x^2$.

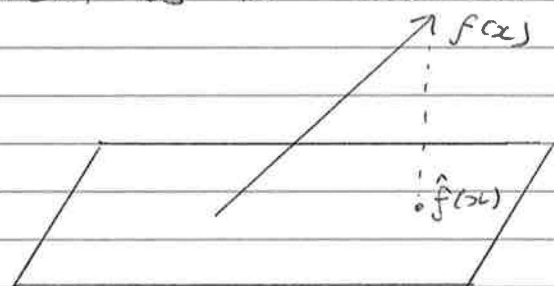
$$W = \{ c_1 P_1 + c_2 P_2 + c_3 P_3 ; \; c_1, c_2, c_3 \in \mathbb{R} \}$$

Note $W$ is closed under addition and scalar multiplication.

<u>Exercise</u> Is $P(x) = -2 + 5x + 3x^2$ in $W$?

Yes! $c_1 = -2$ and $c_2 = 5$ and $c_3 = 3$

Now let's ask the following question. What is the "nearest" element of $W$ closest to $f(x) = \sin(x)$?




$W \equiv \operatorname{span}\{1, x, x^2\}$

Using best approximation theorem, $\hat{f} \equiv$ orthogonal projection of $f$ onto $W$.

Let's define some inner product ( E.g. $\langle f, g \rangle = \int_0^1 f(x)\, g(x)\, dx$ )

Recall orthogonality relation: $f - \hat{f} \perp z \quad z \in W$

$$f - \hat{f} \perp P_1 = 0 \implies \langle f - \hat{f}, P_1 \rangle = 0$$
$$f - \hat{f} \perp P_2 = 0 \implies \langle f - \hat{f}, P_2 \rangle = 0$$
$$f - \hat{f} \perp P_3 = 0 \implies \langle f - \hat{f}, P_3 \rangle = 0$$