Lecture 29

Algorithms for the SVD and bipartite matrices

- Eigenvalues of the matrix [0 A; A' 0] give us the SVD

- This is an instance of a 'bipartite matrix' that arises more often than you might think.

- We'll see a few interesting properties of these matrices, including what happens when you run the Lanczos procedure.

Q:on Precond ✓

HW ✓

The SVD

A: m×n

$A = U \Sigma V^T$

U: m×m orthogonal $U^T U = I$

V: n×n $V^T V = I$

Σ: m×n, diag., non-neg.

By convention $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m \geq 0$

(Sorted in decreasing order)

How to Compute U, Σ, V?

· View 1:

$(\sigma_i\text{'s})^2$ are Eigenvalues of $A^T A$.

Why?

$A^T A = V \Sigma^T U^T U \Sigma V^T$

$= V \Sigma^T \Sigma V^T$

$= V \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix} V^T$

⟺ Eig. val. decomp of $A^T A$.

Given V, Σ

We know

$A = U \Sigma V^T$

$AV = U \Sigma V^T V$

$AV = U \Sigma \xrightarrow{\text{One Col}} A v_i = u_i \cdot \sigma_i$

$A V \Sigma^+$ will give U's associated w/ $\sigma_i \neq 0$.

(all, ↓ just get U_i via)

$\Sigma^+ = \begin{cases} 1/\sigma_i & \sigma_i \neq 0 \\ 0 & \text{else.} \end{cases}$

$u_i = A v_i / \|A v_i\|$.

Not the preferred way to get σ_i, v_i, u_i.

Why?

$K(A^T A) = \|A^T A\|_2 \cdot \|(A^T A)^{-1}\|_2$

$= \sigma_1^2 \cdot \sigma_{\min}^{-2}$

$= \left(\frac{\sigma_1}{\sigma_{\min}}\right)^2 = K(A)^2$.

View 2:

$B = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$

Wielandt Matrix

B: (m+n)×(m+n)

B is Symmetric.

B has a Eigendecomposition.

The λ's of B are

$\pm \sigma_i$ of A along w/ lots of zeros!

Why B?

$B^2 = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$

$= \begin{bmatrix} A^T A & 0 \\ 0 & A A^T \end{bmatrix}$.

Result: the Eigenvalues of B are $\pm \sigma_i$ along with (m+n)−2n zero λ's

Also if

$B z = \sigma_i z$

→ Then if

$z = \begin{bmatrix} x \\ y \end{bmatrix}$

we have

$B z = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} Ay \\ A^T x \end{bmatrix} = \sigma_i \begin{bmatrix} x \\ y \end{bmatrix}$

So $A y = \sigma_i x$ is a sing. value/vector pair.

Note 1:

$B \begin{bmatrix} x \\ -y \end{bmatrix} = \begin{bmatrix} -Ay \\ A^T x \end{bmatrix} = \begin{bmatrix} -\sigma_i x \\ \sigma_i y \end{bmatrix} = -\sigma_i \begin{bmatrix} x \\ -y \end{bmatrix}$

And

$\begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} x \\ -y \end{bmatrix} = x^T x - y^T y$.

This shows if we have $B z = \sigma_i z$ then there is another vect z' s.t. $B z' = -\sigma_i z'$

Proof:

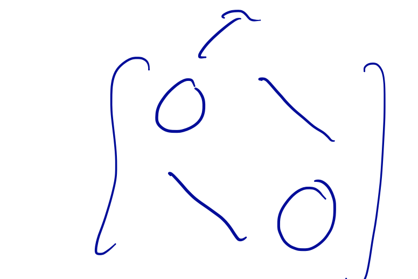
$B = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$

$= \begin{bmatrix} 0 & U \Sigma V^T \\ V \Sigma^T U^T & 0 \end{bmatrix}$

$= \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} 0 & \Sigma \\ \Sigma^T & 0 \end{bmatrix} \begin{bmatrix} U^T & 0 \\ 0 & V^T \end{bmatrix}$

↑

"λ's?"???

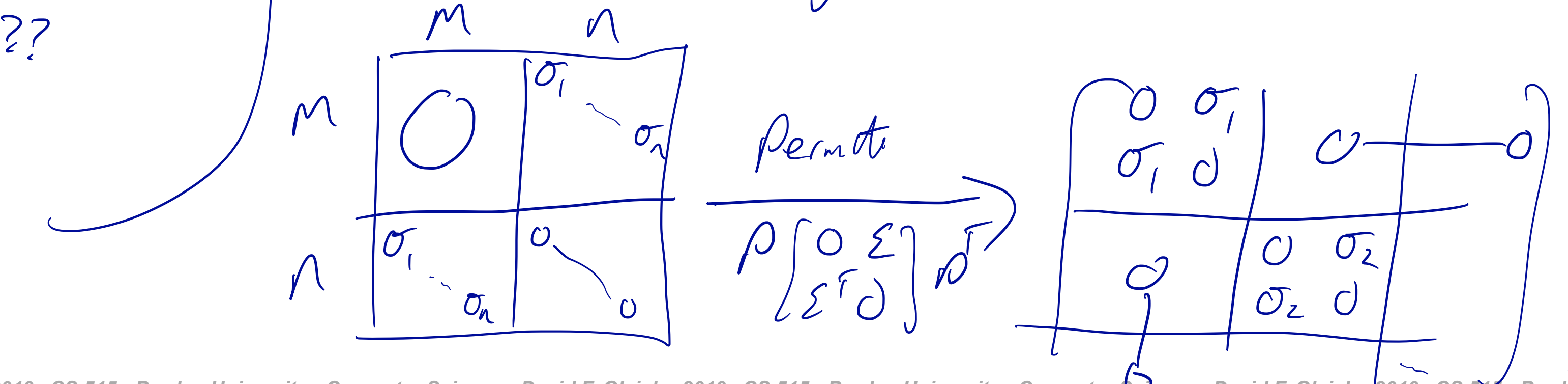


→ If A was 1×1, then we have

$\begin{bmatrix} 0 & \sigma \\ \sigma & 0 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sigma & 0 \\ 0 & -\sigma \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

$(\sigma = |A_{11}|)$.

Now, note that $\begin{bmatrix} 0 & \Sigma \\ \Sigma^T & 0 \end{bmatrix}$ Can be permuted to a block 2×2 diag matrix


m
n
m
n
Permute
$P \begin{bmatrix} 0 & \Sigma \\ \Sigma^T & 0 \end{bmatrix} P^T$

The idea of Algs:

We want to compute λ's of B (and Eigenvectors) but we don't want to even make B!

(*) We don't build B! (*)

(*)

Example:

Compute $\sigma_{\max}$ via Power method on B.

Given z_0: (m+n)-length vect.

for i = 1 to ...

$z_i = \rho_i B z_{i-1}$

$(\rho_i = \frac{1}{\|B z_{i-1}\|})$

$\xrightarrow{\text{Alg w/ A}}$

Given z_0: get (x_0, y_0)

for i = 1 to ...

$x_i = \rho_i A y_{i-1}$; $\rho_i = \frac{1}{\|A y_{i-1}\|_2}$

$y_i = \gamma_i A^T x_{i-1}$, $\gamma_i = \frac{1}{\|A^T x_{i-1}\|_2}$.

Let $z_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$, then $B z_0 = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} A y_0 \\ A^T x_0 \end{bmatrix} \begin{matrix} \to x_1 \\ \to y_1 \end{matrix}$

Step 0 of algs. for all eigenvalues

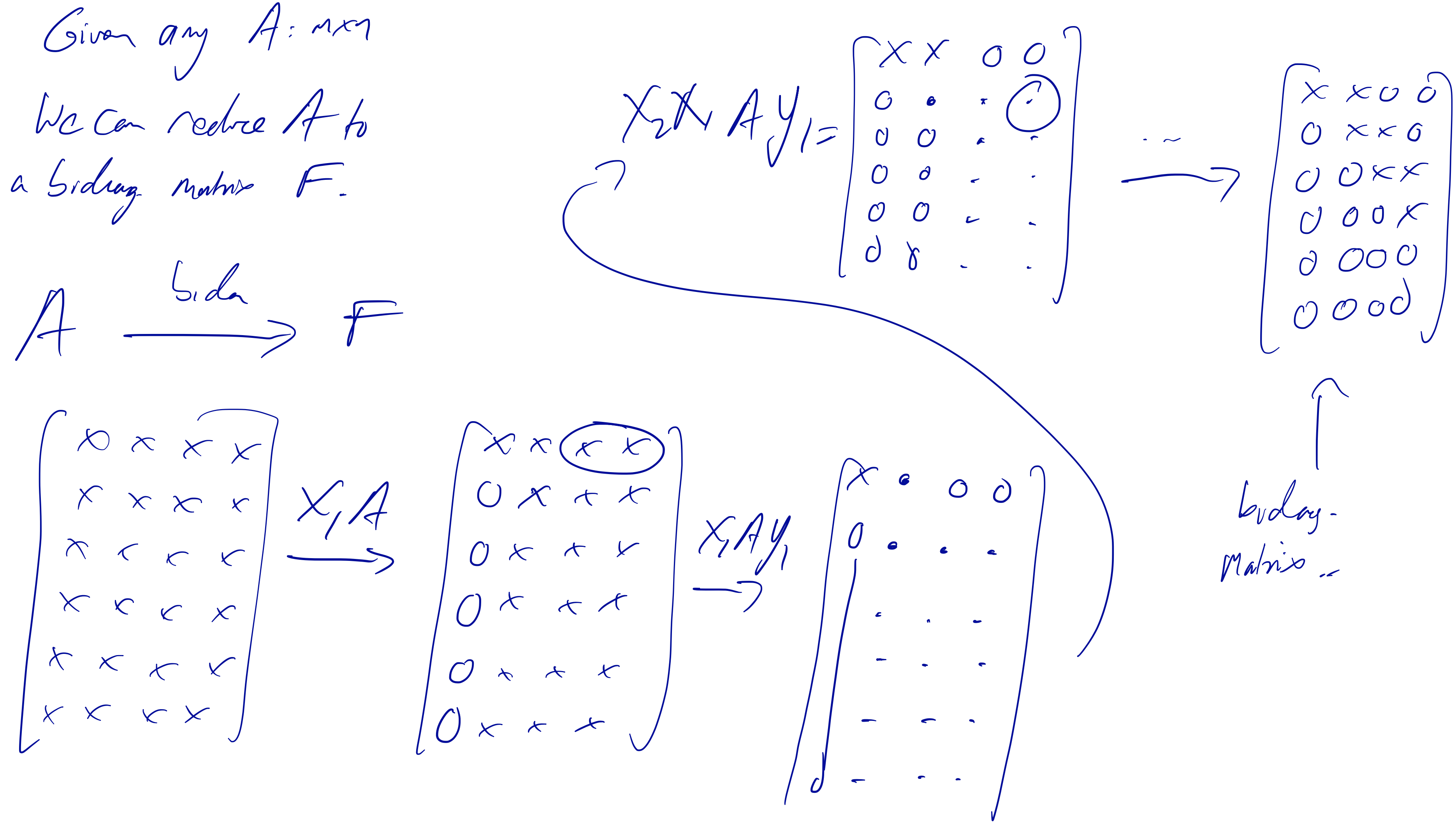
⇒ reduce to tridiag form. (Ortho.)

If $A = U \Sigma V^T$, and X, Y are ortho. then XAY has the same matrix Σ.

Given any A: m×n

We can reduce A to a bidiag matrix F.

$A \xrightarrow{\text{bidiag}} F$



So instead of $B = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$

for SVD, we'd do $\begin{bmatrix} 0 & F \\ F^T & 0 \end{bmatrix}$.

Lanczos on B

$B V_k = V_{k+1} T_{k+1}$

$\uparrow$ k+1 by k tridiag.

V_k is ortho.

Alg:

$$T = \begin{bmatrix} \alpha_1 & \beta_2 & & \\ \beta_2 & \alpha_2 & \beta_3 & \\ & \beta_3 & \ddots & \ddots \\ & & \ddots & \end{bmatrix}$$

$$B \underline{v}_i = \beta_i \underline{v}_{i-1} + \alpha_i \underline{v}_0 + \beta_{i+1} \underline{v}_{i+1}$$

If $v_1 = \begin{bmatrix} \underline{x}_1 \\ 0 \end{bmatrix}$

Then (claim)

$$\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \overbrace{\begin{bmatrix} \underline{x}_1 & 0 & \underline{x}_2 & 0 \\ 0 & y_1 & 0 & y_2 & \cdots \end{bmatrix}}^{V} = \underbrace{\begin{bmatrix} x_1 & 0 & x_2 & 0 \\ 0 & y_1 & 0 & y_2 & \cdots \end{bmatrix}}_{V} \begin{bmatrix} 0 & \beta_2 & & \\ \beta_2 & 0 & \beta_3 & \\ & \beta_3 & 0 & \\ & & & \ddots \end{bmatrix}$$

Why this struct?

$$\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ A^T x_1 \end{bmatrix}$$

$\uparrow$ ortho to this!

Why $\alpha_i = 0$?

$\alpha_i = v_i^T A v_i$.

Idea: use struct.

$$A^T \underline{x}_1 = \beta_2 y_1$$

$$A y_1 = \beta_2 \underline{x}_1 + \beta_3 \underline{x}_2$$

$$A^T \underline{x}_2 = \beta_3 y_1 + \beta_4 y_2$$

$$\vdots$$

$$\Downarrow$$

$$A \begin{bmatrix} y_1 & y_2 \end{bmatrix} = \begin{bmatrix} \underline{x}_1 & \underline{x}_2 & \underline{x}_3 \end{bmatrix} \begin{bmatrix} \beta_2 & 0 \\ \beta_3 & \beta_4 \\ 0 & \beta_5 \end{bmatrix}$$