Eigenvalue Algorithms

CS292F

April 28, 2021

Lecture 8

Q orthogonal matrix: square, $Q^TQ = I$

rotations (around an axis $\equiv$ vector)

or

reflection (through a hyperplane)

or

composition of same.

Q, V orthog. $\Rightarrow QV$ orthog.

$$\|Qx\| = \|x\| \quad \forall x$$

orthogonal transforms don't amplify error

Given $A = A^T$ (any symm. real matrix)

Find W and Λ (diagonal)

with

$$A = W \Lambda W^T \quad \left(\text{or } W^T A W = \Lambda\right)$$

and $W^T W = I$

① Dense $A = A^T$.

Idea : Transform A into $W^T A W = \Lambda$ by multiplying A by orthogonal matrices (lots of them) on left & or right.

TWO STEPS:

① $Q^T A Q = T$ (tridiagonal)

$Q^T Q = I$.

② $V^T T V = \Lambda$ (diagonal)

$V^T V = I$

Then $W = QV$ and $W^T A W = \Lambda$

① looks sort of like Gaussian elimination, zero out rows/cols of A one at a time by Householder reflections. WORK $= \Theta(n^3)$

② the iterative part, uses Givens rotations between just 2 rows/cols. "bulge chasing" "toothpaste squeezing" WORK $\sim \Theta(n^2)$ (modulo numerics & convergence)

SPARSE $A = A^T$.

- SPARSE data structures:
  store only nonzero values.
- Usually don't need <u>all</u> the evals.

<u>But</u> DENSE step 1 causes
n^2 fill-in, n^3 work.

<u>NEED A SPARSE ALGORITHM</u>

Can always compute Ax from x.

THE POWER METHOD

$b = \text{random}; \quad x_0 = b / \|b\|$

for $k = 0, 1, 2, \ldots$ $\qquad A = W \Lambda W^T$

$$\left[\begin{array}{l} v = A x_k \\ \beta_k = \|v\| \\ x_{k+1} = v / \beta_k \end{array}\right.$$

Suppose $x_0 = \sum_{i=1}^{n} \alpha_i w_i \quad (\alpha_i = w_i^T x_0)$

$$x_k = \frac{A^k x_0}{\|A^k x_0\|}$$

suppose λ_n is max $|\cdot|$

$$A^k x_0 = \sum_{i=1}^{n} \alpha_i \lambda_i^k w_i$$

$$= \alpha_n \lambda_n^k \left(w_n + \sum_{i=1}^{n-1} \frac{\alpha_i}{\alpha_n} \cdot \left(\frac{\lambda_i}{\lambda_n}\right)^k w_k \right) \to 0$$

now $\|A^k x_0\| = \left(\sum_i (\alpha_i \lambda_i^k)^2 \right)^{1/2} \longrightarrow (\alpha_n \lambda_n^k)$

so

$$x_k \sim \underline{w_n} + \sum_{i=1}^{n-1} (\text{as above})$$

Convergence is controlled by

$$\left|\frac{\lambda_{n-1}}{\lambda_n}\right|^k \quad (\lambda_{n-1} \text{ second largest } |\cdot|)$$

$$\left|1 - \frac{\lambda_n - \lambda_{n-1}}{\lambda_n}\right|^k$$

Depends on "gap" $|\lambda_n - \lambda_{n-1}|$

---

1. convergence depends on gap.
2. need $\alpha_n \neq 0$ (ok at random)
3. Only get w_n.

$K_k(A,b) = \text{span}(b, Ab, A^2b, \dots, A^{k-1}b)$

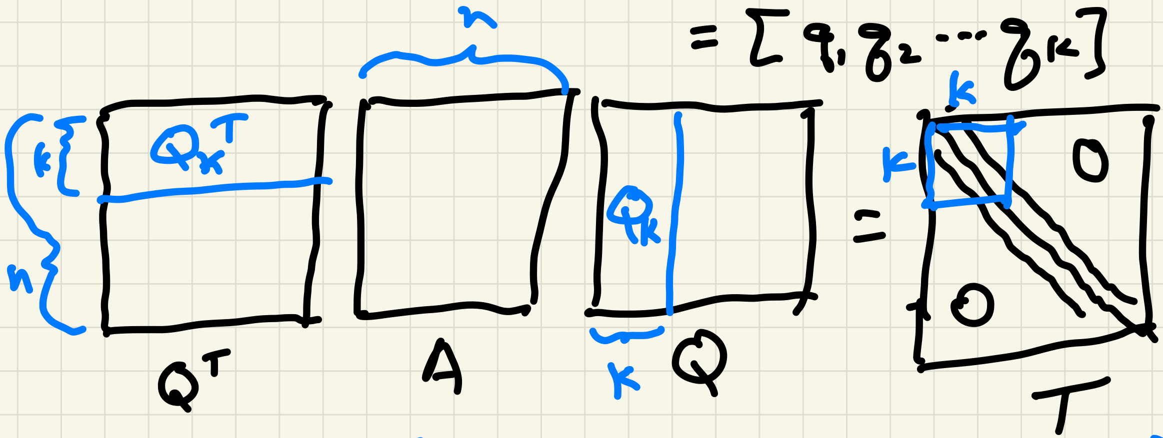
is the k^{th} Krylov subspace of

A and b.

LANCZOS ITERATION

$Q^TAQ = T$. Take $Q_k = Q[:,1:k]$ n x k matrix

$= [q_1\ q_2 \cdots q_k]$



Let $Q_k^T A Q_k = T_k$ (T_k is k by k)

It turns out the cols of Q_k span $K_k(A, q_1)$.

and $Q^TAQ = T$ gives a formula for q_{k+1} in terms of q_{k-1}, q_k, and Aq_k.

$$T = \begin{bmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & 0 \\ & \beta_2 & \alpha_3 & & \\ 0 & & \beta_3 & \ddots & \beta_{n-1} \\ & & & & \alpha_n \end{bmatrix},$$

$$Q^T A Q = T$$

$$Q^T A q_k = k^{\text{th}} \text{ col of } T = \begin{bmatrix} 0 \\ 0 \\ \beta_{k-1} \\ \alpha_k \\ \beta_k \\ 0 \\ 0 \end{bmatrix}$$

t_k

and $A q_k = Q Q^T A q_k = Q t_k$

$$= \beta_{k-1} q_{k-1} + \alpha_k q_k + \beta_k q_{k+1}$$

$$Aq_k = QQ^TAq_k = Qt_k$$

$$= \beta_{k-1} q_{k-1} + \alpha_k q_k + \beta_k q_{k+1}$$

LANCZOS ITERATION

$b = \text{random}$; $q_1 = b/\|b\|$; $q_0 = 0$; $\beta_0 = 0$

for $k = 1, 2, 3, \ldots$

$\quad v = Aq_k$

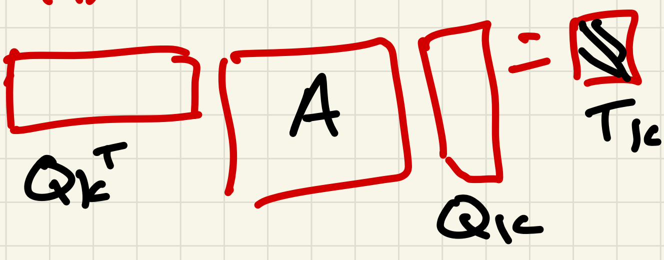
$\quad \alpha_k = q_k^T v$ (because $q_k^T A q_k = T(k,k)$ $= \alpha_k$)

$\quad v = v - \beta_{k-1} q_{k-1} - \alpha_k q_k$

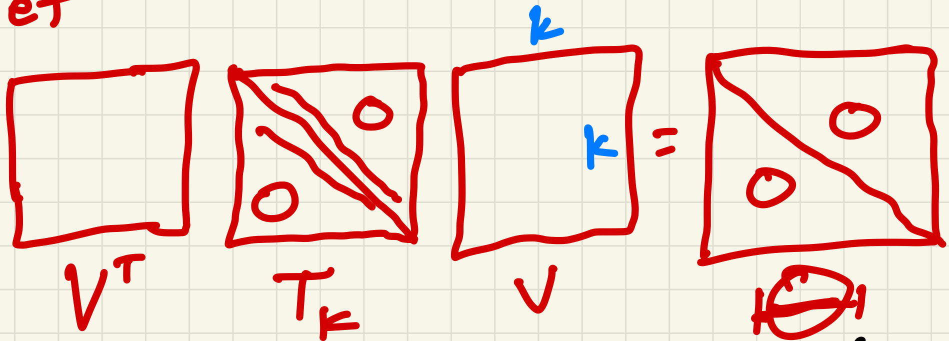
$\quad \beta_k = \|v\|$ (because $\|q_{k+1}\| = 1$)

$\quad q_{k+1} = v / \beta_k$

WE HAVE

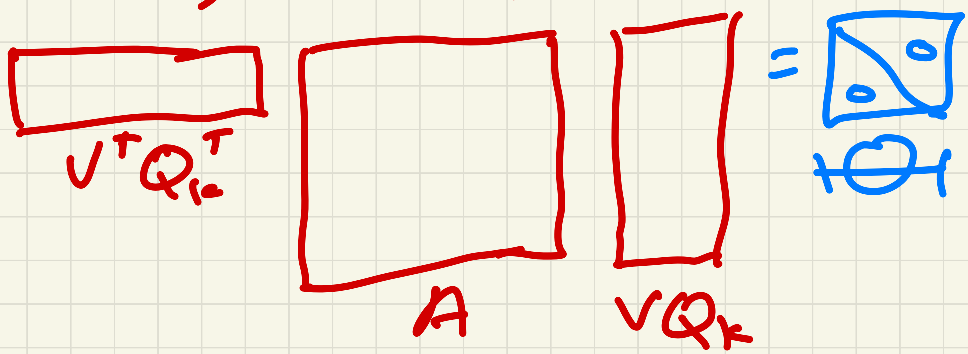


Do k steps of iteration. Then use phase 2 of the dense algorithm to get



Eigenvalues of T_k.

$\Theta = \operatorname{diag}(\theta_1, \theta_2, \ldots, \theta_k)$



EVALS OF A = EVALS OF T. $\neq$ EVALS OF Θ.

The $\theta_1 \ldots \theta_k$ are called "Ritz values"

<u>Theorem:</u> Ritz values are approximations to eigenvalues of A. (not a subset exactly)

THEOREM: Every Ritz value is close to an eval of A.

In particular,

$$\forall_i \exists_j \, |\theta_i - \lambda_j| \leq \beta_k$$ at step k of Lanczos.

$1 \leq i \leq k \quad 1 \leq j \leq n$

THEOREMISH:

The Ritz values usually converge to the extreme eigenvalues first.