Orthogonal Iteration

- A generalization of the power method for computing higher-dimensional invariant subspaces

Algorithm: Method of Orthogonal Iteration
input: $\mathbf{A} \in \mathbb{C}^{n \times n}$, $r \in \{1, \ldots, n\}$, $\mathbf{Q}^{(0)} \in \mathbb{C}^{n \times r}$ semi-unitary
for $k = 1, 2, \ldots$ (until a termination criterion is satisfied)
  $\mathbf{Z}^{(k)} = \mathbf{A}\mathbf{Q}^{(k-1)}$
  Find (thin) QR decomposition of $\mathbf{Z}^{(k)}$: $\mathbf{Q}^{(k)}\mathbf{R}^{(k)} = \mathbf{Z}^{(k)}$, $\mathbf{Q}^{(k)} \in \mathbb{C}^{n \times r}$
  Find the eigenvalues $\lambda_1^{(k)}, \ldots, \lambda_r^{(k)}$ of $(\mathbf{Q}^{(k)})^H \mathbf{A}\mathbf{Q}^{(k)}$
end
output: $\mathbf{Q}^{(k)}$, $\lambda_1^{(k)}, \ldots, \lambda_r^{(k)}$

1st column of $Q^{(k)}$

- $\mathbf{Q}^{(k)}\mathbf{e}_1$ is the vector generated by the power method starting from $\mathbf{v}^{(0)} = \mathbf{Q}^{(0)}\mathbf{e}_1$

- When $r = 1$, the algorithm reduces to the power method

$Q^{(k)}$ is obtained by normalizing $Z^{(k)}$

Analysis of Orthogonal Iteration

Recall the Schur decomposition of $\mathbf{A} \in \mathbb{C}^{n\times n}$

$$\mathbf{U}^H\mathbf{A}\mathbf{U} = \mathbf{T}, \quad \mathbf{U} = \begin{bmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_n\end{bmatrix} \text{ unitary, } \mathbf{T} \text{ uppertriangular, } t_{ii} = \lambda_i$$

$= \mathrm{span}\{u_1, \cdots, u_i\}$

Fact: Let $\mathbf{U}_i = \begin{bmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_i\end{bmatrix}$, $i = 1, \ldots, n$. Then, $\mathcal{R}(\mathbf{U}_i)$ is an invariant subspace for $\mathbf{A}$ and the eigenvalues of $\mathbf{U}_i^H\mathbf{A}\mathbf{U}_i$ are $\lambda_1, \ldots, \lambda_i$

$$U^H A U = T \iff AU = UT \iff Au_i = Ut_i$$

$$\iff Au_i = [u_1 \cdots u_n] \begin{bmatrix} t_{1i} \\ \vdots \\ t_{ii} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = t_{1i}u_1 + \cdots + t_{ii}u_i \quad \forall i = 1, \ldots, n$$

Let $q = \alpha_1 u_1 + \cdots + \alpha_i u_i \in \mathcal{R}(U_i)$

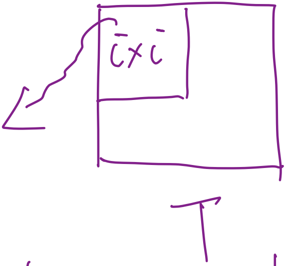
$$Aq = \alpha_1 Au_1 + \cdots + \alpha_i Au_i$$
$$= \alpha_1(t_{11}u_1) + \cdots + \alpha_i(t_{1i}u_1 + \cdots + t_{ii}u_i)$$
$$\in \mathcal{R}(U_i) \longleftarrow \text{invariant subspace for } A$$

Analysis of Orthogonal Iteration (cont'd)

$$U_i^H A U_i = \begin{bmatrix} u_1^H \\ \vdots \\ u_i^H \end{bmatrix} \begin{bmatrix} A u_1 & \cdots & A u_i \end{bmatrix}$$

$$= \begin{bmatrix} u_1^H \\ u_2^H \\ \vdots \\ u_i^H \end{bmatrix} \left[\, t_{11} u_1 \;\middle|\; t_{12} u_1 + t_{22} u_2 \;\middle|\; \cdots \;\middle|\; t_{1i} u_1 + t_{2i} u_2 + \cdots + t_{ii} u_i \,\right]$$

$$= \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1i} \\ 0 & t_{22} & \cdots & t_{2i} \\ & & \ddots & \vdots \\ & & & t_{ii} \end{bmatrix}$$



eigenvalues of $U_i^H A U_i$: $t_{11} = \lambda_1, \ldots, t_{ii} = \lambda_i$

Analysis of Orthogonal Iteration (cont'd)

Suppose the eigenvalues of $\mathbf{A}$ are ordered as

$$|\lambda_1| \geq \cdots \geq |\lambda_r| > |\lambda_{r+1}| \geq \cdots \geq |\lambda_n|, \quad t_{ii} = \lambda_i$$

Partition $\mathbf{U}$ and $\mathbf{T}$ as

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_r & \mathbf{U}_\beta \end{bmatrix}, \quad \mathbf{U}_r \in \mathbb{C}^{n \times r},\ \mathbf{U}_\beta \in \mathbb{C}^{n \times (n-r)}$$

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{0} & \mathbf{T}_{22} \end{bmatrix}, \quad \mathbf{T}_{11} \in \mathbb{C}^{r \times r},\ \mathbf{T}_{22} \in \mathbb{C}^{(n-r) \times (n-r)}$$

With $|\lambda_r| > |\lambda_{r+1}|$, $D_r(\mathbf{A}) := \mathcal{R}(\mathbf{U}_r)$ is called the dominant invariant subspace, which is the unique invariant subspace associated with the eigenvalues $\lambda_1, \ldots, \lambda_r$ of $\mathbf{A}$

Convergence: With proper assumptions,[1]

$$\text{dist}(D_r(\mathbf{A}), \mathcal{R}(\mathbf{Q}^{(k)})) = \mathcal{O}(|\frac{\lambda_{r+1}}{\lambda_r}|^k) \to 0 \text{ as } k \to \infty$$

$$|\lambda_i^{(k)} - \lambda_i| = \mathcal{O}(|\frac{\lambda_{i+1}}{\lambda_i}|^k), \quad i = 1, \ldots, r$$

may be 1

With certain acceleration schemes, $\frac{\lambda_{i+1}}{\lambda_i} \to \frac{\lambda_{r+1}}{\lambda_i} \in [0,1)$

[1] See Theorem 7.3.1 of textbook for details

Orthogonal Iteration (cont'd)

Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ with Schur Decomposition $\mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{T}$, where $\mathbf{U} = \left[\mathbf{u}_1, \ldots, \mathbf{u}_n\right]$ is unitary and $\mathbf{T}$ is uppertriangular with $t_{ii} = \lambda_i$ s.t.

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|, \quad t_{ii} = \lambda_i$$

Let $\mathbf{Q}^{(k)} = \begin{bmatrix} \mathbf{q}_1^{(k)} & \cdots & \mathbf{q}_n^{(k)} \end{bmatrix}$ be generated by the method of orthogonal iteration with $r = n$

It can be shown that with a proper $\mathbf{Q}^{(0)}$,

$$\mathrm{dist}(\mathrm{span}\{\mathbf{q}_1^{(k)}, \ldots, \mathbf{q}_i^{(k)}\}, \mathrm{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_i\}) \to 0 \text{ as } k \to \infty, \quad \forall i = 1, \ldots, n$$

This implies that $\mathbf{T}^k = (\mathbf{Q}^{(k)})^H \mathbf{A} \mathbf{Q}^{(k)}$ converges to an upper triangular matrix, so that the algorithm leads to a Schur decomposition

Orthogonal Iteration (cont'd)

$T^{(k)} = (Q^{(k)})^H A Q^{(k)}$

Compute $\mathbf{T}^k$ more efficiently via its predecessor $\mathbf{T}^{(k-1)}$

QR decomposition of $T^{(k-1)}$

$$\mathbf{T}^{(k-1)} = (\mathbf{Q}^{(k-1)})^H \mathbf{A}\mathbf{Q}^{(k-1)} = (\mathbf{Q}^{(k-1)})^H \mathbf{Z}^{(k)} = \underbrace{(\mathbf{Q}^{(k-1)})^H \mathbf{Q}^{(k)}}_{\text{unitary}} \mathbf{R}^{(k)}$$

$= I$

$$\begin{aligned}\mathbf{T}^{(k)} &= (\mathbf{Q}^{(k)})^H \mathbf{A}\mathbf{Q}^{(k)} = (\mathbf{Q}^{(k)})^H \mathbf{A}\mathbf{Q}^{(k-1)} \cdot (\mathbf{Q}^{(k-1)})^H \mathbf{Q}^{(k)} \\ &= (\mathbf{Q}^{(k)})^H \underbrace{\mathbf{Z}^{(k)}}_{=Q^{(k)}R^{(k)}} \cdot (\mathbf{Q}^{(k-1)})^H \mathbf{Q}^{(k)} = \mathbf{R}^{(k)} (\mathbf{Q}^{(k-1)})^H \mathbf{Q}^{(k)}\end{aligned}$$

This suggests that we may find $\mathbf{T}^{(k)}$ by computing the QR decomposition of $\mathbf{T}^{(k-1)}$ and then multiplying the factors in reverse order

The QR Algorithm/QR Iteration

The above computation of $\mathbf{T}^{(k)}$ motivates the QR algorithm

Algorithm: QR algorithm
input: $\mathbf{A} \in \mathbb{C}^{n \times n}$
$\mathbf{A}^{(0)} = \mathbf{A}$
for $k = 1, 2, \ldots$ (until a termination criterion is satisfied)
  Find QR decomposition of $\mathbf{A}^{(k-1)}$: $\mathbf{Q}^{(k)}\mathbf{R}^{(k)} = \mathbf{A}^{(k-1)}$
  $\mathbf{A}^{(k)} = \mathbf{R}^{(k)}\mathbf{Q}^{(k)}$
end
output: $\mathbf{A}^{(k)}$

$A^{(k-1)}$

similarity transformation

$$A^{(k)} = R^{(k)} Q^{(k)} = (Q^{(k)})^{-1} \boxed{Q^{(k)} R^{(k)}} Q^{(k)}$$

- $\mathbf{A}^{(k)}$ $\forall k$ are similar matrices and thus have the same set of eigenvalues
- If the Schur decomposition of $\mathbf{A}$ is $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^H$, then under some mild assumptions, $\mathbf{A}^{(k)}$ converges to $\mathbf{T}$
  - The diagonal elements of $\mathbf{A}^{(k)}$ for a sufficiently large k would give all the eigenvalues of $\mathbf{A}$
- Complexity of each iteration: $\mathcal{O}(n^3)$
- Improved algorithms can be found in Sections 7.4 and 7.5 of textbook, including the practical QR algorithm (same main idea)

Matrix Computations
Chapter 4: Eigenvalues, Eigenvectors, and Eigendecomposition

Section 4.5 Power method for PageRank

Jie Lu
ShanghaiTech University

Case Study: PageRank

An algorithm used by Google to rank the pages of a search result [1]

More important webpages are likely to receive more links from other websites

Determine the importance of each webpage based on the quality and quantity of links pointing to it

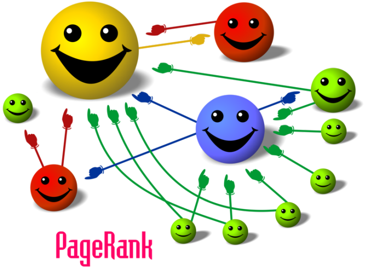


Figure: PageRank. Source: Wikipedia

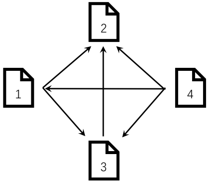
[1] K. Bryan and L. Tanya, "The 25, 000, 000, 000 eigenvector: The linear algebra behind Google," *SIAM Review*, vol. 48, no. 3, pp. 569–581, 2006.

PageRank Model

Let v_i be the importance score of page $i = 1, \ldots, n$, $\mathcal{L}_i$ be the set of pages containing a link to page i, and c_j be the number of outgoing links from page j

$$\sum_{j \in \mathcal{L}_i} \frac{v_j}{c_j} = v_i, \quad \forall i = 1, \ldots, n$$

Example:



$$\begin{aligned} \mathcal{L}_1 &= \{4\} & c_1 &= 2 \\ \mathcal{L}_2 &= \{1,3,4\} & c_2 &= 0 \\ \mathcal{L}_3 &= \{1,4\} & c_3 &= 1 \\ \mathcal{L}_4 &= \emptyset & c_4 &= 3 \end{aligned}$$

$$\begin{bmatrix} 0 & 0 & 0 & \frac{1}{3} \\ \frac{1}{2} & 0 & 1 & \frac{1}{3} \\ \frac{1}{2} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}$$

Notation and Definitions

Notation: For any $x, y \in \mathbb{R}^n$,

- $\mathbf{x} \geq \mathbf{y}$ means that $x_i \geq y_i$ for all i element-wise
- $\mathbf{x} > \mathbf{y}$ means that $x_i > y_i$ for all i
- $\mathbf{x} \not\geq \mathbf{y}$ means that $\mathbf{x} \geq \mathbf{y}$ does not hold
- The same notation applies to matrices

Definitions:

- $\mathbf{x}$ is said to be non-negative if $\mathbf{x} \geq \mathbf{0}$, and non-positive if $-\mathbf{x} \geq \mathbf{0}$
- $\mathbf{x}$ is said to be positive if $\mathbf{x} > \mathbf{0}$, and negative if $-\mathbf{x} > \mathbf{0}$
- The same definitions apply to matrices
- A square matrix $\mathbf{A}$ is said to be column-stochastic if $\mathbf{A} \geq \mathbf{0}$ and $\mathbf{A}^T\mathbf{1} = \mathbf{1}$
    - Each column $\mathbf{a}_i$ of column-stochastic $\mathbf{A}$ satisfies $\mathbf{a}_i^T\mathbf{1} = \sum_{j=1}^n a_{ji} = 1$

PageRank Problem

$A \geq 0$

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a matrix s.t. $a_{ij} = 1/c_j$ if $j \in \mathcal{L}_i$ and $a_{ij} = 0$ if $j \notin \mathcal{L}_i$

Problem: Find a non-negative $\mathbf{v}$ s.t. $\mathbf{Av} = \mathbf{v}$

- $\mathbf{A}$ is extremely large and sparse, so we choose the power method

Questions:

- Does a solution to $\mathbf{Av} = \mathbf{v}$ always exist? Or, is $\lambda = 1$ always an eigenvalue of $\mathbf{A}$?
- Does $\mathbf{Av} = \mathbf{v}$ have a non-negative solution? Or, is there a non-negative eigenvector associated with $\lambda = 1$?
- Is the solution to $\mathbf{Av} = \mathbf{v}$ unique? Or, would there exist more than one eigenvector associated with $\lambda = 1$?
  - A unique solution is desired for PageRank

algebraic multiplicity of $\lambda = 1$ is 1 and no other eigenvalues have modulus 1

- Is $\lambda = 1$ the only eigenvalue that is the largest in modulus?
  - Required by the power method

PageRank Matrix Properties

Observation: In PageRank, $\mathbf{A}$ is column-stochastic if all pages have outgoing links

Properties: Let $\mathbf{A}$ be column-stochastic. Then,

- $\lambda = 1$ is an eigenvalue of $\mathbf{A}$
- $|\lambda| \leq 1$ for any eigenvalue λ of $\mathbf{A}$

Implications: There exists a solution to $\mathbf{A}\mathbf{v} = \mathbf{v}$ and $\lambda = 1$ is an eigenvalue with the largest modulus

Remaining questions: We still don't know

- whether $\mathbf{v} \geq \mathbf{0}$ or not
- whether $\lambda = 1$ is the *only* eigenvalue that has the largest modulus (i.e., whether its algebraic multiplicity is 1 and no other distinct eigenvalues have modulus 1)

We resort to *non-negative matrix theory* to find the answers

Non-Negative Matrix Theory

Theorem (Perron-Frobenius)

Let **A** *be a* *positive* *square matrix. There exists an eigenvalue* ρ *of* **A** *s.t.*

- ρ *is real and* $\rho > 0$
- $|\lambda| < \rho$ *for any eigenvalue* λ *of* **A** *with* $\lambda \neq \rho$
- *There exists a positive eigenvector associated with* ρ
- *The algebraic multiplicity of* ρ *is* 1 *(so the geometric multiplicity of* ρ *is also* 1*)*

Theorem (more general matrix, weaker result)

Let **A** *be a* *non-negative* *square matrix. There exists an eigenvalue* ρ *of* **A** *s.t.*

- ρ *is real and* $\rho \geq 0$
- $|\lambda| \leq \rho$ *for any eigenvalue* λ *of* **A**
- *There exists a non-negative eigenvector associated with* ρ

Modified PageRank Model

From the theorem for non-negative matrices, there exists a non-negative solution to $\mathbf{Av} = \mathbf{v}$, but we don't know whether there exists another solution $\mathbf{v}'$ and whether $\mathbf{v}' \ngeq \mathbf{0}$

For PageRank, we actually consider a modified version of $\mathbf{A}$

$$\tilde{\mathbf{A}} = (1-\beta)\mathbf{A} + \beta \begin{bmatrix} 1/n & \dots & 1/n \\ \vdots & & \vdots \\ 1/n & \dots & 1/n \end{bmatrix}$$

and column stochastic

where $0 < \beta < 1$ (typical value is $\beta = 0.15$), so that $\tilde{\mathbf{A}}$ is positive

From the Perron-Frobenius Theorem,

- $\lambda = 1$ is the only eigenvalue that has the largest modulus
- There exists only one eigenvector associated with $\lambda = 1$, either positive or negative

  dim eigenspace $= 1$
- Therefore, the power method can work

Matrix Computations
Chapter 4: Eigenvalues, Eigenvectors, and Eigendecomposition

Section 4.6 More on Variational Characterizations of Eigenvalues

Jie Lu
ShanghaiTech University

Courant-Fischer Min-Max Theorem (Revisit)

For $\mathbf{A} \in \mathbb{H}^{n\times n}$, let $\lambda_k(\mathbf{A})$ denote the kth largest eigenvalue of $\mathbf{A}$, i.e.,

$$\lambda_{\min}(\mathbf{A}) := \lambda_n(\mathbf{A}) \leq \cdots \leq \lambda_1(\mathbf{A}) =: \lambda_{\max}(\mathbf{A})$$

For simplicity, we may also write $\lambda_{\min} := \lambda_n \leq \cdots \leq \lambda_1 =: \lambda_{\max}$

Theorem

For any $\mathbf{A} \in \mathbb{H}^{n\times n}$ and $k = 1, \ldots, n$,

$$\begin{aligned}\lambda_k(\mathbf{A}) &= \max_{\substack{\mathcal{S}\subseteq\mathbb{C}^n:\\ \dim(\mathcal{S})=k}} \ \min_{\substack{\mathbf{y}\in\mathcal{S},\\ \mathbf{y}\neq\mathbf{0}}} \frac{\mathbf{y}^H\mathbf{A}\mathbf{y}}{\mathbf{y}^H\mathbf{y}} \\ &= \min_{\substack{\mathcal{S}\subseteq\mathbb{C}^n:\\ \dim(\mathcal{S})=n-k+1}} \ \max_{\substack{\mathbf{y}\in\mathcal{S},\\ \mathbf{y}\neq\mathbf{0}}} \frac{\mathbf{y}^H\mathbf{A}\mathbf{y}}{\mathbf{y}^H\mathbf{y}}\end{aligned}$$

$R_{\mathbf{A}}(\mathbf{y}) = \frac{\mathbf{y}^H\mathbf{A}\mathbf{y}}{\mathbf{y}^H\mathbf{y}}$, $\mathbf{y} \neq \mathbf{0}$ is the Rayleigh–Ritz quotient, or Rayleigh quotient

This section focuses on variational characterizations of eigenvalues of real symmetric matrices $(\mathbb{S}^n)$

Rayleigh-Ritz Theorem

A special case of Courant-Fischer Min-Max Theorem (let $k=1$ and $k=n$)

Theorem (Rayleigh-Ritz)

For any $\mathbf{A} \in \mathbb{S}^n$,

$$\lambda_{\min}\|\mathbf{x}\|_2^2 \le \mathbf{x}^T\mathbf{A}\mathbf{x} \le \lambda_{\max}\|\mathbf{x}\|_2^2$$

where the equalities can be attained when $\mathbf{x}$ *is an eigenvector associated with* $\lambda_{\min}$ *and* $\lambda_{\max}$*, respectively*

- Even without Courant-Fischer Min-Max Theorem, we may prove this using eigendecomposition $\mathbf{A} = \mathbf{V}\Lambda\mathbf{V}^T$, $\mathbf{V}$ real orthogonal

$$x^TAx = x^TV\Lambda V^Tx = (V^Tx)^T\Lambda(\underbrace{V^Tx}_{=y}) = y^T\Lambda y$$

$$= \sum_{i=1}^n \lambda_i y_i^2 \le \lambda_{max}\|y\|_2^2 = \lambda_{max}\|V^Tx\|_2^2 = \lambda_{max}\|x\|_2^2$$

Let $x = v_1 \rightsquigarrow \lambda_{max}$ $\quad x^TAx = v_1^TAv_1 = \lambda_{max}\|v_1\|_2^2$

More Results from Courant-Fischer

Let $\mathbf{A}, \mathbf{B} \in \mathbb{S}^n$, $\mathbf{z} \in \mathbb{R}^n$

- (Weyl) $\lambda_k(\mathbf{A}) + \lambda_n(\mathbf{B}) \le \lambda_k(\mathbf{A}+\mathbf{B}) \le \lambda_k(\mathbf{A}) + \lambda_1(\mathbf{B})$, $k = 1, \ldots, n$

$$\lambda_k(A+B) = \min_{\substack{S \subseteq \mathbb{R}^n \\ \dim(S) = n-k+1}} \max_{\substack{x \in S \\ \|x\|_2 = 1}} x^T (A+B) x$$

$$\le \min_{\substack{S \subseteq \mathbb{R}^n \\ \dim(S) = n-k+1}} \left(\max_{\substack{x \in S \\ \|x\|_2 = 1}} x^T A x + \max_{\boxed{\substack{x \in S \\ \|x\|_2 = 1}}} x^T B x \right)$$

$$x \in \mathbb{R}^n, \ \|x\|_2 = 1$$

$$\le$$

$$= \min_{\substack{S \subseteq \mathbb{R}^n \\ \dim(S) = n-k+1}} \max_{\substack{x \in S \\ \|x\|_2 = 1}} x^T A x + \min_{\substack{S \subseteq \mathbb{R}^n \\ \dim(S) = n-k+1}} \lambda_1(B)$$

$$= \lambda_k(A) + \lambda_1(B)$$

More Results from Courant-Fischer (cont'd)

Let $\mathbf{A}, \mathbf{B} \in \mathbb{S}^n$, $\mathbf{z} \in \mathbb{R}^n$

- (Interlacing) $\lambda_{k+1}(\mathbf{A}) \leq \lambda_k(\mathbf{A} \pm \mathbf{z}\mathbf{z}^T) \leq \lambda_{k-1}(\mathbf{A})$ for proper k

$$\lambda_k(A \pm zz^T) = \min_{\substack{S \subseteq \mathbb{R}^n \\ \dim(S) = n-k+1}} \max_{\substack{x \in S \\ \|x\|_2 = 1}} x^T(A \pm zz^T)x$$

$$\geq \min_{\substack{S \subseteq \mathbb{R}^n \\ \dim(S) = n-k+1}} \max_{\substack{x \in S \cap \operatorname{span}\{z\}^\perp \\ \|x\|_2 = 1}} (x^TAx \pm \underbrace{x^Tzz^Tx}_{=0})$$

Note that $\dim(S \cap \operatorname{span}\{z\}^\perp) = \overset{n-k+1}{\dim(S)} + \overset{n-1}{\dim(\operatorname{span}\{z\}^\perp)} - \underbrace{\dim(S + \operatorname{span}\{z\}^\perp)}_{\leq n} \geq n-k$

It follows that

$$\lambda_k(A \pm zz^T) \geq \min_{\substack{S' \subseteq \mathbb{R}^n : \dim(S) = r \\ r \in [n-k, n]}} \max_{\substack{x \in S' \\ \|x\|_2 = 1}} x^TAx \geq \lambda_{k+1}(A)$$

$r = n-k \rightsquigarrow \lambda_{k+1}(A)$ $\quad r = n \rightsquigarrow \lambda_1(A)$

$r = n-k+1 \rightsquigarrow \lambda_k(A)$