New

Eigenvalues: first-order sensitivity

Given A, compute eigenvalues & eigenvectors:

$$\underset{\uparrow\ \text{input}}{A}v = \overset{\text{output}}{\lambda}\, \underset{\text{output}}{v}$$

$$(\underset{\text{perturbed input}}{A+\delta A})(v+\delta v) = \underbrace{(\lambda+\delta\lambda)(v+\delta v)}_{\text{perturbed outputs}}$$

Suppose A is diagonalizable, and has distinct eigenvalues:

$$A = \underbrace{\begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} - & w_1^* & - \\ & \vdots & \\ - & w_n^* & - \end{bmatrix}}_{V^{-1}}$$

w_j - left eigenvector of A

$$\underbrace{\begin{bmatrix} - & w_1^* & - \\ & \vdots & \\ - & w_n^* & - \end{bmatrix}}_{V^{-1}} A = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \underbrace{\begin{bmatrix} - & w_1^* & - \\ & \vdots & \\ - & w_n^* & - \end{bmatrix}}_{V^{-1}}$$

$$\Rightarrow \quad w_j^* A = \lambda_j w_j^*$$

Dropping 2^{nd} order terms in (1), we have

$$\underbrace{\cancel{Av} + \delta A v + A\delta v = \cancel{\lambda v}}_{Av=\lambda v} + \lambda \delta v + \delta\lambda v$$

multiply by w^*: $\quad w^*\delta A v + \cancel{w^* A}\delta v = \lambda \cancel{w^*}\delta v + \delta\lambda w^* v \qquad w^*A = \lambda w^*$

$$\Rightarrow \quad \delta\lambda = \frac{w^*\delta A v}{w^* v}$$

$$\boxed{\left|\frac{\delta\lambda}{\lambda}\right| \leq \underbrace{\frac{\|w\|_2\|v\|_2}{|w^*v|}}_{\text{"Wikinson's Condition \#"}} \frac{\|\delta A\|}{|\lambda|}}$$

Normal Matrices

Normal matrices have orthogonal eigenvectors:

$$\begin{bmatrix} - & v_1^* & - \\ & \vdots & \\ - & v_n^* & - \end{bmatrix} = V^* = V^{-1} = \begin{bmatrix} - & w_1^* & - \\ & \vdots & \\ - & w_n^* & - \end{bmatrix}$$

$$\Rightarrow \quad v_i = w_i$$

left and right eigenvectors are the same!

Then, $\dfrac{\|w\|\|v\|}{|w^*v|} = \dfrac{\|v\|^2}{\|v\|^2} = 1$ so eigenvalues of normal matrices are perfectly well-cond:

$$\Rightarrow \left|\frac{\delta\lambda}{\lambda}\right| \leq \frac{\|\delta A\|}{|\lambda|}$$

Real symmetric (complex Hermitian), Orthogonal (unitary) and skew-symmetric (skew-Hermitian) are all important examples of normal matrices.

<u>Example</u>

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1+\delta \end{bmatrix}, \qquad \delta > 0$$

Eigenvalues: $\lambda_1 = 1$, $\lambda_2 = 1+\delta$

Eigenvectors: $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} \delta^{-1} \\ 1 \end{pmatrix}$

$$w_1 = \begin{pmatrix} -\delta \\ 1 \end{pmatrix} \quad w_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Wilkinson's Condition #: $\kappa_j = \dfrac{\|w_j\| \|v_j\|}{|w_j^* v_j|}$

$$\kappa_1 = \frac{\sqrt{1+\delta^2}}{\delta} \qquad \kappa_2 = \sqrt{1+1/\delta^2} = \frac{\sqrt{1+\delta^2}}{\delta}$$

$$\kappa_1, \kappa_2 \to \infty \quad \text{as} \quad \delta \to 0!$$

Let's look at perturbations to A when

$\delta = 0$, i.e., when condition # is "infinite."

$$A_1^{(\varepsilon)} = \begin{bmatrix} 1 & 1 \\ \varepsilon & 1 \end{bmatrix} \Rightarrow \lambda_\pm^{(\varepsilon)} = 1 \pm \sqrt{\varepsilon}$$

so $|\lambda^{(\varepsilon)} - \lambda| = \sqrt{\varepsilon} = O(\sqrt{\delta A_1})$

where $A_1^{(\varepsilon)} = \underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}_{A_1} + \underbrace{\begin{bmatrix} 0 & 0 \\ \varepsilon & 0 \end{bmatrix}}_{\delta A_1}$.

Perturbation of λ grows proportional to $\sqrt{\delta A}$ instead of δA, hence why the condition # is undefined. This happens b/c $\lambda^{(\varepsilon)}$ is continuous but not differentiable at $\varepsilon = 0$.

<u>Pseudospectra</u> (See "Pseudospectra Gateway")

A powerful tool for analyzing the sensitivity of eigenvalues and related phenom.

is the ε-pseudospectra of A:

(a) $\lambda_\varepsilon(A) = \{z : \|(A - zI)v\| < \varepsilon \text{ for some } v \in \mathbb{R}^n\}$

Idea: Instead of requiring $Av = \lambda v$, look for nearby values of z that are "almost" eigenvectors of A.

What makes pseudospectra so useful is the following equivalent characterizations:

(b) $\lambda_\varepsilon(A) = \{z \in \lambda(A + \delta A), \text{ where } \|\delta A\| < \varepsilon\}$

(c) $\lambda_\varepsilon(A) = \{z : \|(A - zI)^{-1}\| > 1/\varepsilon\}$

They allow us to understand how far eigenvalues can "travel" under perturbations of size ε (b) by bounding the resolvent norm (c).

For normal matrices: $A = V \Lambda V^*$

$$\|(A - zI)^{-1}\| = \|V(\Lambda - zI)^{-1} V^*\|$$
$$= \|(\Lambda - zI)^{-1}\|$$

$$= \left(\min_{1 \le j \le n} |\lambda_j - z|\right)^{-1}$$

$$\Rightarrow \lambda_\varepsilon(A) = \lambda(A) + \Delta_\varepsilon$$

open ball of radius ε

Similarly for non-normal matrices

$$\lambda_\varepsilon(A) \subset \lambda(A) + \Delta_{\varepsilon \kappa(V)}$$

condition # of eigenmatrix.

First-order sensitivity of eigenvectors

$\hat{\lambda}, \hat{x}$ approx. λ, x where $Ax = \lambda x$

$A^* = A$ symm.

(normalization $\|\hat{x}\| = \|x\| = 1$)

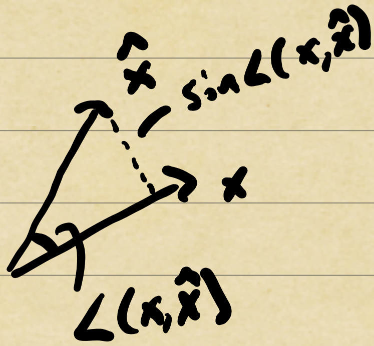
Residual $r = A\hat{x} - \hat{\lambda}\hat{x}$

"Sin θ theorem" (Davis-Kahan)

$$\sin \angle(x, \hat{x}) \le \frac{\|r\|}{\min_{\lambda_j \neq \lambda} |\hat{\lambda} - \lambda_j|}$$

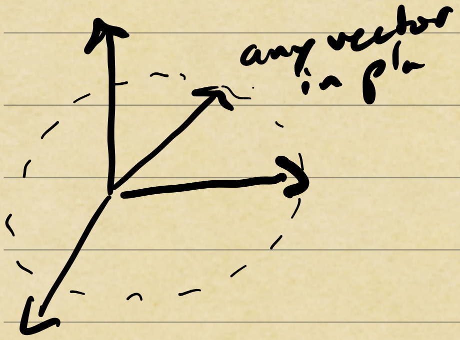
other eigenvalues of A

$Ax_j = \lambda_j x_j$



When $\lambda, \hat{\lambda}$ is close to rest of spectrum relative to $\|r\|$ ("small gap"), eigenvectors direction can change a lot!

Intuition: When $\lambda_k = \lambda_j$ for some $k \neq j$ there is a whole plane of eigenvectors for $\lambda = \lambda_k = \lambda_j$. Small disturbances break the plane into just two eigendirections, and any perturbation that "passes through" multiple eigenvalues allows disc. changes in eigenvectors ($\kappa_{A \to v} = \infty$)



Note that the invariant <u>subspace</u> – the span of eigenvectors – associated w/a cluster of eigenvalues may be entirely <u>well-conditioned</u>, even though the individual eigenvectors are <u>not</u>.

⇒ see G. W. ("Pete") Stewart's work on invariant subspaces for more.