Eigenvalue Solvers

Introduction to Sensitivity and Accuracy

Eigenvalue problems

Before we compute eigenvalues and eigenvectors numerically, we must understand what we can and cannot compute (accurately) or should not compute.

We may want $Ax = \lambda x$ for single eigenpair, for example with λ as small as possible. Say minimum energy level.

In many cases we want $Ax_i = \lambda_i x_i$ for $i = 1 \ldots M$ (M large), where λ_i are smallest M eigenvalues.

- It may be important that we do not skip any eigenvalues.
- We may want the invariant subspace accurately.
- We may want every eigenvector accurately.

Usefulness of Computed Results

In general we need to consider the accuracy of a computed answer, without knowing the exact answer.

This involves the sensitivity of the result we want to compute.

If some result is very sensitive to small changes in the problem, it may be impossible to compute exactly. In other cases results may be computable but at very high price, for example, an algorithm may convergence very slowly.

Sometimes it is better to compute related but less sensitive result.

Sensitivity of an Eigenvalue

Sensitivity of eigenvalues to perturbations in the matrix.

Different eigenvalues or eigenvectors of a matrix are not equally sensitive to perturbations of the matrix.

Let $Ax = \lambda x$ and $y^* A = \lambda y^*$, where $\|x\| = \|y\| = 1$.

Consider $(A + E)(x + e) = (\lambda + \varepsilon)(x + e)$ and drop second order terms.

$$Ax + Ae + Ex \cong \lambda x + \lambda e + \varepsilon x \Leftrightarrow Ae + Ex \cong \lambda e + \varepsilon x$$

$$y^* Ae + y^* Ex = \lambda y^* e + y^* Ex \cong \lambda y^* e + \varepsilon y^* x \quad \Leftrightarrow$$

$$y^* Ex \cong \varepsilon y^* x \Rightarrow \varepsilon \cong \frac{y^* Ex}{y^* x} \Rightarrow |\varepsilon| \leq \frac{\|E\|}{y^* x}$$

Condition number of simple eigenvalue: $\kappa(\lambda) = \left| y^* x \right|^{-1}$

Sensitivity of an Eigenvalue

For symmetric/Hermitian matrix, right and left eigenvectors are the same. So, eigenvalues are inherently well-conditioned.

More generally, eigenvalues are well conditioned for normal matrices, but eigenvalues of nonnormal matrices need not be well conditioned.

Nonnormal matrices may not have a full set of eigenvectors. The algebraic multiplicity, the multiplicity of λ as a root of $\det(A - \lambda I) = 0$, is not equal to the geometric multiplicity, $\dim \operatorname{null}(A - \lambda I)$. In that case we can consider conditioning of the invariant subspace associated with a Jordan block.

Sensitivity of an Eigenvalue

If μ is an eigenvalue of $A+E$, then $\lambda \in \mathcal{L}(A)$ exists:

$$|\mu - \lambda| \leq \left\| XEX^{-1} \right\| \leq \kappa(X)\|E\|$$

where X is the matrix of eigenvectors of A and
$\kappa(X) \equiv \|X\|\left\|X^{-1}\right\|$ is a condition number (consistent norm).

A useful backward error result is given by the residual.

Let $r = Ax - \lambda x$ and $\|x\| = 1$.

Then there exists a perturbation E with $\|E\| \leq \|r\|$ such that

$$(A+E)x = \lambda x.$$

Proof: Take $E = -rx^*$.

Sensitivity of Eigenvectors

Consider $A = \begin{pmatrix} 1 & 0 \\ 0 & 1+\varepsilon_1 \end{pmatrix} \quad \rightarrow \quad x_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \; x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$

Consider perturbation E with $\|E\| = \varepsilon_1 + \varepsilon_2$, ε_2 arb. small.

Enough to give A any eigenvectors not equal to x_1 and x_2.

Let $E = E_1 + E_2$ and $\hat{X} = \begin{bmatrix} \hat{x}_1 & \hat{x}_2 \end{bmatrix}$ (unitary)

Let $E_1 = \begin{pmatrix} 0 & 0 \\ 0 & -\varepsilon_1 \end{pmatrix}$ and $E_2 = \hat{X} \begin{pmatrix} 0 & 0 \\ 0 & \varepsilon_2 \end{pmatrix} \hat{X}^*$.

$A + E_1 = I$ (all nonzero vectors are eigenvectors)

$$A + E = I + \hat{X} \begin{pmatrix} 0 & 0 \\ 0 & \varepsilon_2 \end{pmatrix} \hat{X}^* = \hat{X} \begin{pmatrix} 1 & 0 \\ 0 & 1+\varepsilon_2 \end{pmatrix} \hat{X}^*$$

Eigenpair Perturbation

General perturbation result is not simple

Let (λ, x) be simple eigenpair of A, $\begin{pmatrix} x & X \end{pmatrix}$ be nonsingular,

and $\begin{pmatrix} y & Y \end{pmatrix}^* A \begin{pmatrix} x & X \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & M \end{pmatrix}$, where $\begin{pmatrix} y & Y \end{pmatrix}^* = \begin{pmatrix} x & X \end{pmatrix}^{-1}$.

Let $\tilde{A} = A + E$, $\begin{pmatrix} y & Y \end{pmatrix}^* E \begin{pmatrix} x & X \end{pmatrix} = \begin{pmatrix} \varphi_{11} & f_{12}^* \\ f_{21} & F_{22} \end{pmatrix}$, and

$\operatorname{sep}(\lambda, \mathrm{M}) = \left\| (\lambda I - M)^{-1} \right\|^{-1}$. (consistent norm)

Eigenpair Perturbation

If $4\|f_{21}\|\|f_{12}^*\| < \left(\operatorname{sep}(\lambda, M) - |\varphi_{11}| - \|F_{22}\|\right)^2$, then $\left(\tilde{\lambda}, \tilde{x}\right) = (\lambda + \varphi, x + Xp)$, where

$$\|p\| < \frac{2\|f_{21}\|}{\left(\operatorname{sep}(\lambda, M) - |\varphi_{11}| - \|F_{22}\|\right)},$$

$$\left\|p - (\lambda I - M)^{-1} f_{21}\right\| < \frac{2\|p\|^2 \left\|f_{12}^*\right\|}{\operatorname{sep}(\lambda, M) - |\varphi_{11}| - \|F_{22}\|},$$

$$\left|\varphi - y^* E x\right| \leq \|p\|\|f_{21}\|$$

If Y orthonormal, $\sin \measuredangle(x, \tilde{x}) \leq \dfrac{\|E\|_2}{\operatorname{sep}\left(\tilde{\lambda}, M\right)}$ (2-norm).

Eigenpair Perturbation

Hermitian/symmetric case:

$$\operatorname{sep}(\lambda, M) = \min\left\{|\lambda - \mu| : \mu \in \mathcal{L}(M)\right\}$$

$$\sin \measuredangle(x, \tilde{x}) \lesssim \frac{\|E\|_2}{\min_\mu |\lambda - \mu|}$$

Distance to nearest other eigenvalue determines angle.

$$\left|\varphi - y^* E x\right| \leq \|p\| \|f_{21}\|$$

$$\sin \measuredangle(x, \tilde{x}) \leq \frac{\|E\|_2}{\operatorname{sep}(\lambda + \varphi, M)} \text{ (2-norm).}$$

Accuracy

A small residual means a small backward error. The result is exact for a slightly perturbed problem. For numerical computations that is usually the best you can get.

However, if the problem is ill-conditioned, the answer may not be accurate.
$Ax - \lambda x = r$ implies $(A + E)x = \lambda x$ with $\|E\| \leq \|r\|$.

So, if λ too close to eigenvalues in $\mathcal{L}(M)$ relative to $\|r\|$, then (λ, x) may be completely wrong. Note that if some eigenvalues in $\mathcal{L}(M)$ are ill-conditioned, $\operatorname{sep}(\lambda, M)$ may be small even if $\min\{|\lambda - \mu| : \mu \in \mathcal{L}(M)\}$ is not.

Computing Invariant Subspace

So, we should avoid defining (sub)problems that are sensitive.

For example, if we want to compute eigenpairs up to some eigenvalue (energy level). We can expect problems if the selection includes some eigenvalues from a tight cluster but not all.

If the distance between two eigenvalues is very small relative to norm of the matrix, they are numerically identical: multiple eigenvalue.

Even if we can distinguish them their eigenvectors may be very ill-conditioned, and so on.

Computing Invariant Subspace

More complicated cases can also arise. Especially for nonnormal problems.

If an eigenvalue is very sensitive, a small perturbation may move the eigenvalue near or in another cluster.

This means it may not be a good idea to compute an invariant subspace that has small canonical angles with the invariant subspace corresponding to other eigenvalues.

This type of ill-conditioning not as easy to spot as close eigenvalues.

General Perturbation Result

Let $\begin{pmatrix} X_1 & X_2 \end{pmatrix}$ be unitary and $\begin{pmatrix} X_1 & X_2 \end{pmatrix}^* A \begin{pmatrix} X_1 & X_2 \end{pmatrix} = \begin{pmatrix} L_1 & H \\ G & L_2 \end{pmatrix}$.

The $\text{range}(X_1)$ is invariant subspace iff $G = X_2^* A X_1 = 0$.

How near is $\text{range}(X_1)$ to invariant subspace for G small?

Let P be matrix such that

$$\hat{X}_1 = (X_1 + X_2 P)\left(I + P^* P\right)^{-1/2}$$

$$\hat{X}_2 = \left(X_2 - X_1 P^*\right)\left(I + P P^*\right)^{-1/2}$$

where $\text{range}\left(\hat{X}_1\right)$ is invariant subspace.

General Perturbation Result

Define $T(P) \equiv PL_1 + L_2P$. P under following condition.

$$\delta = \operatorname{sep}(L_1, L_2) \equiv \inf_{\|P\|=1} \|T(P)\| > 0,\ \eta = \|H\|,\ \gamma = \|G\|.$$

If $4\gamma\eta < \delta^2$ then unique P exists such that $\|P\| < 2\dfrac{\gamma}{\delta}$.

$\hat{X}_1, \hat{X}_2$ generate invar. subspaces with representation $\hat{L}_1$, $\hat{L}_2$.

$$\hat{L}_1 = \left(I + P^*P\right)^{1/2} \left(L_1 + HP\right)\left(I + P^*P\right)^{-1/2}$$

$$\hat{L}_2 = \left(I + PP^*\right)^{-1/2} \left(L_2 - PH\right)\left(I + PP^*\right)^{1/2}$$

General Backward Error Result

Let $S = X_1^* A - L_1 X_1^*$. Let $L_1 = X_1^* A X_1$ and $L_2 = X_2^* A X_2$.

$R = AX_1 - X_1 L_1$ and $S = X_1^* A - L_1 X_1^*$

If $\dfrac{\|R\|\|S\|}{\operatorname{sep}(L_1, L_2)} < \dfrac{1}{4}$

then simple invariant subspace, $\operatorname{range}\left(\hat{X}_1\right)$, such that

$$\left\|\tan\Theta(X_1, X_2)\right\| < 2\frac{\|R\|}{\operatorname{sep}(L_1, L_2)}$$

References

1. G.W. Stewart, *Matrix Algorithms, Volume II: Eigensystems*, SIAM
2. G.W. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press.