Perturbation Theory for Eigenvalue Problems

Nico van der Aa

October 19th 2005

Overview of talks

- Erwin Vondenhoff (21-09-2005)
  A Brief Tour of Eigenproblems
- Nico van der Aa (19-10-2005)
  Perturbation analysis
- Peter in 't Panhuis (9-11-2005)
  Direct methods
- Luiza Bondar (23-11-2005)
  The power method
- Mark van Kraaij (7-12-2005)
  Krylov subspace methods
- Willem Dijkstra (...)
  Krylov subspace methods 2

Outline of my talk

Goal

My goal is to illustrate ways to deal with sensitivity theory of eigenvalues and eigenvectors.

Way

By means of examples I would like to illustrate the theorems.

Assumptions

There are no special structures present in the matrices under consideration. They are general complex valued matrices.

Recap on eigenvalue problems

Definition of eigenvalue problems

$$\boldsymbol{A}\boldsymbol{X} - \boldsymbol{X}\boldsymbol{\Lambda} = \boldsymbol{0}, \qquad \boldsymbol{Y}^*\boldsymbol{A} - \boldsymbol{\Lambda}\boldsymbol{Y}^* = \boldsymbol{0}$$

with * the complex conjugate transposed and

$$\underbrace{\boldsymbol{X} = \begin{bmatrix} | & | & & | \\ \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_n \\ | & | & & | \end{bmatrix}}_{\text{right eigenvectors}}, \quad \underbrace{\boldsymbol{Y}^* = \begin{bmatrix} - & \boldsymbol{y}_1 & - \\ - & \boldsymbol{y}_2 & - \\ & \vdots & \\ - & \boldsymbol{y}_n & - \end{bmatrix}}_{\text{left eigenvectors}}, \quad \underbrace{\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}}_{\text{eigenvalues}}$$

The left-eigenvectors are chosen such that

$$\boldsymbol{Y}^*\boldsymbol{X} = \boldsymbol{I}$$

Bauer-Fike Theorem

Theorem

Given are λ an eigenvalue and $\boldsymbol{X}$ the matrix consisting of eigenvectors of matrix $\boldsymbol{A}$. Let μ be an eigenvalue of matrix $\boldsymbol{A}+\boldsymbol{E} \in \mathbb{C}^{n\times n}$, then

$$\min_{\lambda\in\sigma(\boldsymbol{A})} |\lambda-\mu| \leq \underbrace{\|\boldsymbol{X}\|_p \|\boldsymbol{X}^{-1}\|_p}_{K_p(\boldsymbol{X})} \|\boldsymbol{E}\|_p \qquad (\dagger)$$

where $\|.\|_p$ is any matrix p-norm and $K_p(\boldsymbol{X})$ is called the condition number of the eigenvalue problem for matrix $\boldsymbol{A}$.

Proof

The proof can be found in many textbooks.

- *Numerical Methods for Large Eigenvalue Problems*
  Yousef Saad
- *Numerical Mathematics*
  A. Quarteroni, R. Sacco, F. Saleri

Bauer-Fike Theorem (2)

Example

$$\boldsymbol{A} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}, \qquad \boldsymbol{\Lambda} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}, \qquad \boldsymbol{X} = \begin{bmatrix} 1 & \frac{1}{2}\sqrt{2} \\ 0 & \frac{1}{2}\sqrt{2} \end{bmatrix}.$$

$$\boldsymbol{E} = \begin{bmatrix} 0 & 0 \\ 10^{-4} & 0 \end{bmatrix}, \quad K_2(\boldsymbol{X}) \approx 2.41, \quad \|\boldsymbol{E}\|_2 = 10^{-4}$$

The Bauer-Fike theorem states that the eigenvalues can change 2.41×10^{-4}. In this example, they only deviate $1e - 4$.

Remarks

- The Bauer-Fike theorem is an over estimate.
- The Bauer-Fike theorem does not give a direction.

Eigenvalue derivatives - Theory

Suppose that $\boldsymbol{A}$ depends on a parameter p and its eigenvalues are distinct. The derivative of the eigensystem is given by

$$\boldsymbol{A}'(p)\boldsymbol{X}(p) - \boldsymbol{X}(p)\boldsymbol{\Lambda}'(p) = -\boldsymbol{A}(p)\boldsymbol{X}'(p) + \boldsymbol{X}'(p)\boldsymbol{\Lambda}(p).$$

Premultiplication with the left-eigenvectors gives

$$\boldsymbol{Y}^*\boldsymbol{A}'\boldsymbol{X} - \underbrace{\boldsymbol{Y}^*\boldsymbol{X}}_{=\boldsymbol{I}}\boldsymbol{\Lambda}' = \boldsymbol{Y}^*\boldsymbol{A}\boldsymbol{X}' + \boldsymbol{Y}^*\boldsymbol{X}'\boldsymbol{\Lambda}.$$

Introduce $\boldsymbol{X}' = \boldsymbol{X}\boldsymbol{C}$. This is allowed since for distinct eigenvalues the eigenvectors form a basis of $\mathbb{C}^n$. Then,

$$\boldsymbol{Y}^*\boldsymbol{A}'\boldsymbol{X} - \boldsymbol{\Lambda}' = -\underbrace{\boldsymbol{Y}^*\boldsymbol{A}\boldsymbol{X}}_{=\boldsymbol{\Lambda}}\boldsymbol{C} + \underbrace{\boldsymbol{Y}^*\boldsymbol{X}}_{=\boldsymbol{I}}\boldsymbol{C}\boldsymbol{\Lambda}.$$

Written out in components, the eigenvalue derivatives is given by

$$\lambda_k' = \boldsymbol{y}_k^*\boldsymbol{A}'\boldsymbol{x}_k$$

Eigenvalue derivatives - Example

Example definition

$$\boldsymbol{A} = \begin{bmatrix} p & 1 \\ 1 & -p \end{bmatrix}, \quad \boldsymbol{A}' = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

In this case, the eigenvalues can be computed analytically

$$\boldsymbol{\Lambda} = \begin{bmatrix} -\sqrt{p^2+1} & 0 \\ 0 & \sqrt{p^2+1} \end{bmatrix}, \quad \boldsymbol{\Lambda}' = \begin{bmatrix} -\frac{p}{\sqrt{p^2+1}} & 0 \\ 0 & \frac{p}{\sqrt{p^2+1}} \end{bmatrix}$$

The method for $p = 1$

The following quantities can be computed from the given matrix $\boldsymbol{A}(p)$

$$\boldsymbol{A}(1) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \boldsymbol{\Lambda}(1) = \begin{bmatrix} -\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}, \quad \boldsymbol{X}(1) = \begin{bmatrix} 0.3827 & -0.9239 \\ -0.9239 & -0.3827 \end{bmatrix}, \quad \boldsymbol{Y}^*(1) = \begin{bmatrix} 0.3827 & -0.9239 \\ -0.9239 & -0.3827 \end{bmatrix}$$

The eigenvalue derivatives can be computed by

$$\lambda_1'(1) = \begin{bmatrix} 0.3827 & -0.9239 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0.3827 \\ -0.9239 \end{bmatrix} = -\frac{1}{2}\sqrt{2}$$

$$\lambda_2'(1) = \begin{bmatrix} -0.9239 & -0.3827 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -0.9239 \\ -0.3827 \end{bmatrix} = \frac{1}{2}\sqrt{2}$$

Eigenvector derivatives

Theory

As long as the eigenvalues are distinct, the eigenvectors form a basis of $\mathbb{C}^n$ and therefore the following equation holds:

$$\boldsymbol{Y}^*\boldsymbol{A}'\boldsymbol{X} - \boldsymbol{\Lambda}' = -\boldsymbol{\Lambda}\boldsymbol{C} + \boldsymbol{C}\boldsymbol{\Lambda}.$$

Since

$$(\boldsymbol{\Lambda}\boldsymbol{C} + \boldsymbol{C}\boldsymbol{\Lambda})_{ij} = -\lambda_i c_{ij} + c_{ij}\lambda_j = c_{ij}(\lambda_j - \lambda_i),$$

the off-diagonal entries of $\boldsymbol{C}$ can be determined as follows

$$c_{ij} = \frac{\boldsymbol{y}_i^*\boldsymbol{A}'\boldsymbol{x}_j}{\lambda_j - \lambda_i}, \qquad i \neq j.$$

What about the diagonal entries?

$\Rightarrow$ additional assumption.

Eigenvector derivatives - Normalization

Problem description

An eigenvector is determined uniquely in case of distinct eigenvalues up to a constant.

If matrix $\boldsymbol{A}$ has an eigenvector $\boldsymbol{x}_k$ belonging to eigenvalue λ_k, then $\gamma \boldsymbol{x}_i$ with γ a nonzero constant, is also an eigenvector.

$$\boldsymbol{A}(\gamma \boldsymbol{x}_k) - \lambda_k(\gamma \boldsymbol{x}_k) = \gamma \left(\boldsymbol{A}\boldsymbol{x}_k - \lambda_k \boldsymbol{x}_k\right) = \mathbf{0}$$

Conclusion: there is one degree of freedom to determine the eigenvector itself and therefore also the derivative contains a degree of freedom.

$$(c_k \boldsymbol{x}_k)' = c_k' \boldsymbol{x}_k + c_k \boldsymbol{x}_k'$$

Important: the eigenvector derivative that will be computed is the derivative of this normalized eigenvector!

Eigenvector derivatives - Normalization 2

Solution

A mathematical choice is to set one element of the eigenvector equal to 1 for all p.
How do you choose these constants?

- $\max_{l=1,\ldots,n} |x_{kl}|$;
- $\max_{l=1,\ldots,n} |x_{kl}||y_{kl}|$.

The derivative is computed from the normalized eigenvector.

Remark: the derivative of the element set to 1 for all p is equal to 0 for all p.

Eigenvector derivatives - Normalization 3

Result

Consider only one eigenvector. Its derivative can be expanded as follows:

$$x'_{kl} = \sum_{m=1}^{n} x_{km} c_{ml}.$$

By definition the derivative of the element set to 1 for all p is equal to zero. Therefore,

$$0 = x_{kk} c_{kk} + \sum_{\substack{m=1 \\ m \neq l}}^{n} x_{km} c_{mk} \Rightarrow c_{kk} = -\frac{1}{x_{kk}} \sum_{\substack{m=1 \\ m \neq l}}^{n} x_{km} c_{mk}.$$

Repeating the normalization procedure for all eigenvectors enables the computation of the diagonal entries of $\boldsymbol{C}$.

Finally, the eigenvector derivatives can be computed as follows:

$$\boldsymbol{X}' = \boldsymbol{X}\boldsymbol{C}$$

with $\boldsymbol{X}$ the normalized eigenvector matrix.

Eigenvector derivatives - Example

$$\boldsymbol{A}=\begin{bmatrix}0 & -\frac{ip(-1+p^2)}{1+p^2}\\ \frac{ip(1+p^2)}{-1+p^2} & 0\end{bmatrix},\quad \boldsymbol{A}'=\begin{bmatrix}0 & \frac{i(-1+4p^2+p^4)}{(1+p^2)^2}\\ \frac{i(1+4p^2-p^4)}{(-1+p^2)^2} & 0\end{bmatrix},\quad \boldsymbol{\Lambda}=\begin{bmatrix}-ip & 0\\ 0 & ip\end{bmatrix},\quad \boldsymbol{X}=\begin{bmatrix}1-p^2 & 1-p^2\\ 1+p^2 & -p^2-1\end{bmatrix}$$

Consider the case where $p=2$.

The matrices are given by

$$\boldsymbol{A}=\begin{bmatrix}0 & -\frac{6i}{5}\\ -\frac{10i}{3} & 0\end{bmatrix}$$

$$\boldsymbol{A}'=\begin{bmatrix}0 & -\frac{31i}{25}\\ \frac{i}{9} & 0\end{bmatrix}$$

$$\boldsymbol{X}=\begin{bmatrix}-0.5145 & 0.5145\\ 0.8575 & 0.8575\end{bmatrix}$$

$$\boldsymbol{Y}^*=\begin{bmatrix}-0.9718 & 0.5831\\ 0.9718 & 0.5831\end{bmatrix}$$

The off-diagonal entries of the coefficient matrix $\boldsymbol{C}$ are

$$c_{12}=\frac{\boldsymbol{y}_1^*\boldsymbol{A}'\boldsymbol{x}_2}{\lambda_2-\lambda_1}=-\frac{8}{3}\qquad c_{21}=\frac{\boldsymbol{y}_2^*\boldsymbol{A}'\boldsymbol{x}_1}{\lambda_1-\lambda_2}=-\frac{8}{3}$$

Normalization: for all k and l the following is true $|x_{kl}||y_{kl}|=\frac{1}{2}$. Therefore, choose

$$\boldsymbol{X}=\begin{bmatrix}-\frac{3}{5} & \frac{3}{5}\\ 1 & 1\end{bmatrix}$$

Then the diagonal entries of matrix $\boldsymbol{C}$ become

$$c_{11}=-\frac{x_{22}}{x_{21}}c_{21}=\frac{8}{3}\qquad c_{22}=-\frac{x_{21}}{x_{22}}c_{12}=\frac{8}{3}$$

The eigenvector derivatives can now be computed:

$$\boldsymbol{X}'=\boldsymbol{X}\boldsymbol{C}=\begin{bmatrix}\frac{8}{25} & -\frac{8}{25}\\ 0 & 0\end{bmatrix}$$

Repeated eigenvalues

Problem statement

If repeated eigenvalues occur, that is $\lambda_k = \lambda_l$ for some k and l, then any linear combination of eigenvectors $\boldsymbol{x}_k$ and $\boldsymbol{x}_l$ is also an eigenvector.

To apply the previous theory, we have to make the eigenvectors unique up to a constant multiplier.

Solution procedure

Assume the n known eigenvectors are linearly independent and denote them by $\tilde{\boldsymbol{X}}$. Define

$$\hat{\boldsymbol{X}} = \tilde{\boldsymbol{X}}\boldsymbol{\Gamma} \text{ for some coefficient matrix } \boldsymbol{\Gamma}$$

If the columns of $\boldsymbol{\Gamma}$ can be defined unique up to a constant multiplier, also $\hat{\boldsymbol{X}}$ is uniquely defined up to a constant multiplier.

Repeated eigenvalues - mathematical trick

Computing Γ

Differentiate the eigenvalue system $\boldsymbol{A}\hat{\boldsymbol{X}} = \hat{\boldsymbol{X}}\boldsymbol{\Lambda}$:

$$\boldsymbol{A}'\hat{\boldsymbol{X}} - \hat{\boldsymbol{X}}\boldsymbol{\Lambda}' = -\boldsymbol{A}\hat{\boldsymbol{X}}' + \hat{\boldsymbol{X}}'\boldsymbol{\Lambda}$$

Premultiply with the left-eigenvectors and use the fact that the eigenvalues are repeated

$$\tilde{\boldsymbol{Y}}^*\boldsymbol{A}'\tilde{\boldsymbol{X}}\boldsymbol{\Gamma} - \tilde{\boldsymbol{Y}}^*\tilde{\boldsymbol{X}}\boldsymbol{\Gamma}\boldsymbol{\Lambda}' = -\tilde{\boldsymbol{Y}}^*\underbrace{\left(\boldsymbol{A}\hat{\boldsymbol{X}}' - \hat{\boldsymbol{X}}'\boldsymbol{\Lambda}\right)}_{=(\boldsymbol{A}-\lambda\boldsymbol{I})\hat{\boldsymbol{X}}}$$

Eliminate the right-hand-side

$$\tilde{\boldsymbol{Y}}^*\boldsymbol{A}'\tilde{\boldsymbol{X}}\boldsymbol{\Gamma} - \boldsymbol{\Gamma}\boldsymbol{\Lambda}' = -\underbrace{\tilde{\boldsymbol{Y}}^*(\boldsymbol{A} - \lambda\boldsymbol{I})}_{=0}\hat{\boldsymbol{X}}'$$

Assume that $\lambda_k' \neq \lambda_l'$ for all $k \neq l$, then $\boldsymbol{\Gamma}$ consists of the eigenvectors of matrix $\tilde{\boldsymbol{Y}}^*\boldsymbol{A}'\tilde{\boldsymbol{X}}$ and are determined up to a constant.

Repeated eigenvalues - Example

Computations of the eigenvalues for $p = 2$

Matrix $\boldsymbol{A}$ is constructed from an eigenvector matrix and an eigenvalue matrix with values $\lambda_1 = ip$ and $\lambda_2 = -i(p-4)$. This results in

$$\boldsymbol{A} = \begin{bmatrix} 2i & -\frac{i(-2+p)(-1+p^2)}{1+p^2} \\ -\frac{i(-2+p)(1+p^2)}{-1+p^2} & 2i \end{bmatrix}.$$

For $p = 2$, the eigenvalues become repeated and Matlab gives the following results

$$\boldsymbol{A} = \begin{bmatrix} 2i & 0 \\ 0 & 2i \end{bmatrix}, \quad \boldsymbol{\Lambda} = \begin{bmatrix} 2i & 0 \\ 0 & 2i \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

From the construction of matrix $\boldsymbol{A}$, we know that $\lambda_1' = i$ and $\lambda_2' = -i$, but when we follow the procedure from before, we see that

$$\tilde{\boldsymbol{Y}}^* \boldsymbol{A}' \tilde{\boldsymbol{X}} = \begin{bmatrix} 0 & -0.6i \\ -1.67i & 0 \end{bmatrix} \neq \boldsymbol{\Lambda}'.$$

Now, with the mathematical trick

$$\boldsymbol{\Gamma} = \begin{bmatrix} -0.5145 & 0.5145 \\ 0.8575 & 0.8575 \end{bmatrix}, \quad \hat{\boldsymbol{X}} = \tilde{\boldsymbol{X}}\boldsymbol{\Gamma} = \begin{bmatrix} -0.5145 & 0.5145 \\ 0.8575 & 0.8575 \end{bmatrix}.$$

Repeat the procedure

$$\hat{\boldsymbol{Y}}^* \boldsymbol{A}' \hat{\boldsymbol{X}} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = \boldsymbol{\Lambda}'.$$

Repeated eigenvalues - Extension

Theory

To determine the eigenvector derivatives in the distinct case, the first order derivative of the eigensystem was considered. This does not work since

$$\boldsymbol{Y}^*\boldsymbol{A}'\boldsymbol{X} - \boldsymbol{\Lambda}' = -\boldsymbol{Y}^*\underbrace{(\boldsymbol{A}\boldsymbol{X}' - \boldsymbol{X}'\boldsymbol{\Lambda})}_{=(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{X}'} = \boldsymbol{0}$$

Consider one differentiation higher

$$\boldsymbol{A}''\boldsymbol{X} - \boldsymbol{X}\boldsymbol{\Lambda}'' = -2\boldsymbol{A}'\boldsymbol{X}' + 2\boldsymbol{X}'\boldsymbol{\Lambda}' - \boldsymbol{A}\boldsymbol{X}'' + \boldsymbol{X}''\boldsymbol{\Lambda}$$

Premultiply with the left-eigenvectors and use $\boldsymbol{X}' = \boldsymbol{X}\boldsymbol{C}$, then

$$\boldsymbol{Y}^*\boldsymbol{A}''\boldsymbol{X} - \boldsymbol{\Lambda}'' = -2\underbrace{\boldsymbol{Y}^*\boldsymbol{A}'\boldsymbol{X}}_{=\boldsymbol{\Lambda}'}\boldsymbol{C} + 2\boldsymbol{C}\boldsymbol{\Lambda}' - \underbrace{\boldsymbol{Y}^*(\boldsymbol{A}\boldsymbol{X}'' - \boldsymbol{X}''\boldsymbol{\Lambda})}_{=0}$$

Thus the off-diagonal entries of matrix $\boldsymbol{C}$ is

$$c_{ij} = \frac{y_i^*\boldsymbol{A}''\boldsymbol{x}_j}{2(\lambda_j' - \lambda_i')}, \qquad i \neq j$$

Repeated eigenvalues - Example continued

$$A=\begin{bmatrix} 2i & -\frac{i(-2+p)(-1+p^2)}{1+p^2} \\ -\frac{i(-2+p)(1+p^2)}{-1+p^2} & 2i \end{bmatrix},\quad \Lambda=\begin{bmatrix} -ip & 0 \\ 0 & i(p-4) \end{bmatrix},\quad X=\begin{bmatrix} 1-p^2 & 1-p^2 \\ 1+p^2 & -p^2-1 \end{bmatrix}$$

Consider the case where $p=2$.

The matrices are given by

$$A=\begin{bmatrix} 2i & 0 \\ 0 & 2i \end{bmatrix}$$

$$A'=\begin{bmatrix} 0 & -\frac{3}{5}i \\ -\frac{5}{3}i & 0 \end{bmatrix}$$

$$A''=\begin{bmatrix} 0 & -\frac{16}{25}i; \frac{16}{9}i & 0 \end{bmatrix}$$

$$\hat{X}=\begin{bmatrix} -0.5145 & 0.5145 \\ 0.8575 & 0.8575 \end{bmatrix}$$

The off-diagonal entries of the coefficient matrix C are

$$c_{12}=\frac{y_1^* A'' x_2}{2(\lambda_2'-\lambda_1')}=-\frac{8}{15} \qquad c_{21}=\frac{y_2^* A'' x_1}{2(\lambda_1'-\lambda_2')}=-\frac{8}{15}$$

Normalization: for all k and l the following is true $|x_{kl}||y_{kl}|=\frac{1}{2}$. Therefore, choose

$$\hat{X}=\begin{bmatrix} -\frac{3}{5} & \frac{3}{5} \\ 1 & 1 \end{bmatrix}$$

Then the diagonal entries of matrix C become

$$c_{11}=-\frac{x_{22}}{x_{21}}c_{21}=\frac{8}{15} \qquad c_{22}=-\frac{x_{21}}{x_{22}}c_{12}=\frac{8}{15}$$

The eigenvector derivatives can now be computed:

$$X'=XC=\begin{bmatrix} -\frac{8}{25} & \frac{8}{25} \\ 0 & 0 \end{bmatrix}$$

Conclusions

- Distinct eigenvalues
  - Eigenvalue derivatives can be computed directly from the eigenvectors and the derivative of the original matrix;
  - Eigenvector derivatives can be computed as soon as it is normalized in some mathematical sensible way.
- Repeated eigenvalues
  - A mathematical trick is required to compute the eigenvalue derivatives;
  - To compute the eigenvector derivatives, the second order derivatives of the eigensystem has to be computed.

References

- real-valued matrices
  - Distinct eigenvalues
    - Nelson, R.B., *Simplified Calculation of Eigenvector Derivatives*, AIAA Journal **14**(**9**), September 1976.
  - Repeated eigenvalues
    - Curran, W.C., *Calculation of Eigenvector Derivatives for Structures with Repeated Eigenvalues*, AIAA Journal **26**(**7**), July 1988.
- complex-valued matrices
  - Murthy, D.V. and Haftka, R.T., *Derivatives of Eigenvalues and Eigenvectors of a General Complex Matrix*, International Journal for Numerical Methods in Engineering **26**, pg. 293-311, 1988.

Questions ?