Lecture 10: Eigenvectors and eigenvalues

(*Numerical Recipes*, Chapter 11)

The eigenvalue problem,

$$\mathbf{A}\,\mathbf{x} = \lambda\,\mathbf{x},$$

occurs in many, many contexts:

classical mechanics, quantum mechanics, optics……

Eigenvectors and eigenvalues
(*Numerical Recipes*, Chapter 11)

The textbook solution is obtained from

$(\mathbf{A} - \lambda \mathbf{I})\, \mathbf{x} = 0$, which can only hold for $\mathbf{x} \neq \mathbf{0}$ if

$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$.

This is a polynomial equation for λ of order N

Eigenvectors and eigenvalues

(*Numerical Recipes*, Chapter 11)

This method of solution is:

(1) very slow

(2) inaccurate, in the presence of roundoff error

(3) yields only eigenvalues, not eigenvectors

(4) of no practical value

(except, perhaps in the 2 x 2 case where the solution to the resultant quadratic can be written in closed form)

Review of some basic definitions in linear algebra

A matrix is said to be:

<u>Symmetric</u> if $\mathbf{A} = \mathbf{A}^\mathbf{T}$ (its *transpose*) i.e. if $a_{ij} = a_{ji}$

<u>Hermitian</u> if $\mathbf{A} = \mathbf{A}^\dagger \equiv (\mathbf{A}^T)^*$ i.e. if $a_{ij} = a_{ji}^*$
(its *Hermitian conjugate*)

<u>Orthonormal</u> if $\mathbf{U}^{-1} = \mathbf{U}^T$

<u>Unitary</u> if $\mathbf{U}^{-1} = \mathbf{U}^\dagger$

Review of some basic definitions in linear algebra

All 4 types are *normal*, meaning that they obey the relations

$$\mathbf{A}\,\mathbf{A}^\dagger = \mathbf{A}^\dagger\,\mathbf{A}$$

$$\mathbf{U}\,\mathbf{U}^\dagger = \mathbf{U}^\dagger\,\mathbf{U}$$

Eigenvectors and eigenvalues

Hermitian matrices are of particular interest because they have:

real eigenvalues

orthonormal eigenvectors ($x_i^\dagger x_j = \delta_{ij}$)

(unless some eigenvalues are degenerate, in which case we can always design an orthonormal set)

Diagonalization of Hermitian matrices

Hermitian matrices can be diagonalized according to

Unitary matrix whose column are the eigenvectors

$$\mathbf{A} = \mathbf{U}\,\mathbf{D}\,\mathbf{U}^{-1}$$

Diagonal matrix consisting of the eigenvalues

Diagonalization of Hermitian matrices

Clearly, if $\mathbf{A} = \mathbf{U}\,\mathbf{D}\,\mathbf{U}^{-1}$, then $\mathbf{A}\,\mathbf{U} = \mathbf{U}\,\mathbf{D}$

…..which is simply the eigenproblem (written out N times):

The columns of **U** are the eigenvectors (mutually orthogonal) and the elements of **D** (non-zero only on the diagonal) are the corresponding eigenvalues.

Non-Hermitian matrices

We are often interested in non-symmetric real matrices:

- the eigenvalues are real or come in complex-conjugate pairs
- the eigenvectors are not orthonormal in general
- the eigenvectors may not even span an N-dimensional space

Diagonalization of non-Hermitian matrices

For a non-Hermitian matrix, we can identify two (different) types of eigenvectors

Right hand eigenvectors are column vectors which obey:

$$\mathbf{A}\,\mathbf{x_R} = \lambda\,\mathbf{x_R}$$

Left hand eigenvectors are row vectors which obey:

$$\mathbf{x_L}\,\mathbf{A} = \lambda\,\mathbf{x_L}$$

The eigenvalues are the same (the roots of det $(\mathbf{A} - \lambda\mathbf{I}) = 0$ in both cases) but, in general, the eigenvectors are not

Diagonalization of non-Hermitian matrices

Note: for a Hermitian matrix, the two types of eigenvectors are equivalent, since

$$A\, x_R = \lambda\, x_R$$

since $A = A^\dagger$ since λ is real

$$\Rightarrow\; x_R^\dagger A^\dagger = \lambda^* x_R^\dagger \;\Rightarrow\; x_R^\dagger A = \lambda\, x_R^\dagger$$

so if any vector $\mathbf{x_R}$ is a right-hand eigenvector, then $\mathbf{x_R}^\dagger$ is a left-hand eigenvector

Diagonalization of non-Hermitian matrices

- Let **D** be the diagonal matrix whose elements are the eigenvalues, **D** = diag (λ_i)
- Let $\mathbf{X_R}$ be the matrix whose columns are the right-hand eigenvectors
- Let $\mathbf{X_L}$ be the matrix whose rows are the left-hand eigenvectors
- Then the eigenvalue equations are

$$A\,\mathbf{X_R} = \mathbf{X_R}\,D \qquad \text{and} \qquad \mathbf{X_L}\,A = D\,\mathbf{X_L}$$

Diagonalization of non-Hermitian matrices

- The eigenvalue equations

  $\mathbf{A}\,\mathbf{X_R} = \mathbf{X_R}\,\mathbf{D}$ and $\mathbf{X_L}\,\mathbf{A} = \mathbf{D}\,\mathbf{X_L}$

  imply $\mathbf{X_L} A\,\mathbf{X_R} = \mathbf{X_L}\,\mathbf{X_R}\,\mathbf{D}$ and $\mathbf{X_L}\,\mathbf{A}\,\mathbf{X_R} = \mathbf{D}\,\mathbf{X_L}\,\mathbf{X_R}$

  Hence, $\mathbf{X_L}\,\mathbf{A}\,\mathbf{X_R} = \mathbf{X_L}\,\mathbf{X_R}\,\mathbf{D} = \mathbf{D}\,\mathbf{X_L}\,\mathbf{X_R}$

  ➔ $\mathbf{X_L}\,\mathbf{X_R}$ commutes with $\mathbf{D}$

  ➔ $\mathbf{X_L}\,\mathbf{X_R}$ is also diagonal

  ➔ rows of $\mathbf{X_L}$ are orthogonal to columns of $\mathbf{X_R}$

Diagonalization of non-Hermitian matrices

It follows that any matrix can be diagonalized according to

$\mathbf{D} = \mathbf{X_L}\ \mathbf{A}\ \mathbf{X_R} = \mathbf{X_R}^{-1}\ \mathbf{A}\ \mathbf{X_R}$, or equivalently

$\mathbf{D} = \mathbf{X_L}\ \mathbf{A}\ \mathbf{X_L}^{-1}$

The special feature of a Hermitian matrix is that the $\mathbf{X_L}$ and $\mathbf{X_R}$ are unitary

Solving the eigenvalue equation

The problem is entirely equivalent to figuring out how to diagonalize of A according to

$$\mathbf{A} = \mathbf{X_R}\,\mathbf{D}\,\mathbf{X_R}^{-1}$$

All methods proceed by making a series of similarity transformation

$$\mathbf{A} = \mathbf{P_1}^{-1}\,\mathbf{M_1}\,\mathbf{P_1} = \mathbf{P_1}^{-1}\,\mathbf{P_2}^{-1}\,\mathbf{M_2}\,\mathbf{P_2}\,\mathbf{P_1} = \ldots = \mathbf{X_R}\,\mathbf{D}\,\mathbf{X_R}^{-1}$$

where $\mathbf{M_1}$, $\mathbf{M_2}$ … are successively more nearly diagonal

Solving the eigenvalue equation

There are many routines available:

- EISPACK + Linpack – LAPACK (free)
- NAG, IMSL (expensive)

Before using a totally general technique, consider what you want:

- eigenvalues alone, or eigenvectors as well?

..and how special your case is

- real symmetric & tridiagonal, real symmetric, real non-symmetric, complex Hermitian, complex non-Hermitian

Solution for a real symmetric matrix: the Householder method (*Recipes*, §11.2)

The Householder method makes use of a similarity transform based upon

$$P = \mathrm{I} - 2\, w\, w^T,$$

where $\mathbf{w}$ is a column vector for which $\mathbf{w}^T \mathbf{w} = 1$ and $\mathbf{w}\,\mathbf{w}^T$ is an N x N symmetric matrix

i.e. $(\mathbf{w}\,\mathbf{w}^T)_{ij} = (\mathbf{w}\,\mathbf{w}^T)_{ji} = w_i\, w_j$

The Householder method

The Householder matrix, **P**, clearly has the following properties

P is symmetric (being the difference between 2 symmetric matrices, $\mathbf{I}$ and $2\mathbf{w}\mathbf{w}^\mathbf{T}$)

P is orthogonal:

$$\mathbf{P}^\mathbf{T}\,\mathbf{P} = (\mathbf{I} - 2\,\mathbf{w}\,\mathbf{w}^\mathbf{T})^\mathbf{T}\,(\mathbf{I} - 2\,\mathbf{w}\,\mathbf{w}^\mathbf{T})$$

$$= (\mathbf{I} - 4\,\mathbf{w}\,\mathbf{w}^\mathbf{T} + 4\,\mathbf{w}\,\underbrace{\mathbf{w}^\mathbf{T}\,\mathbf{w}}_{=1}\,\mathbf{w}^\mathbf{T}) = \mathbf{I}$$

The Householder method

Let $\mathbf{a_1}$ be the first column of $\mathbf{A}$

Let $$w = \frac{\mathbf{a_1} - |\mathbf{a_1}|\, \mathbf{e_1}}{|\, \mathbf{a_1} - |\mathbf{a_1}|\, \mathbf{e_1}|}$$

where $\mathbf{e_1} = (1,0,0,0,0\ldots0)^T$

The Householder method

Let's consider the action of $\mathbf{P} = \mathbf{I} - 2\,\mathbf{w}\,\mathbf{w}^{\mathbf{T}}$ upon the first column of $\mathbf{A}$

$$\mathbf{P}\,\mathbf{a_1} = \left\{ \mathbf{I} - \frac{2\,(\mathbf{a_1} - |\mathbf{a_1}|\,\mathbf{e_1})\,(\mathbf{a_1} - |\mathbf{a_1}|\,\mathbf{e_1})^{T}}{|\mathbf{a_1} - |\mathbf{a_1}|\,\mathbf{e_1}|^{2}} \right\} \mathbf{a_1}$$

$$= \mathbf{a_1} - \frac{2\,(\mathbf{a_1} - |\mathbf{a_1}|\,\mathbf{e_1})\,(|\mathbf{a_1}|^{2} - |\mathbf{a_1}|\,\mathbf{e_1}^{T}\mathbf{a_1})}{|\mathbf{a_1}|^{2} - 2\,|\mathbf{a_1}|\,\mathbf{e_1}.\,\mathbf{a_1} + |\mathbf{a_1}|^{2}}$$

$$= |\mathbf{a_1}|\,\mathbf{e_1}$$

➔ $\mathbf{P}$ zeroes out all but the first element of $\mathbf{a_1}$

The Householder method

Suppose **A** is symmetric

$$\mathbf{A} = \begin{pmatrix} x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \end{pmatrix} \quad \Rightarrow \quad \mathbf{P\,A} = \begin{pmatrix} x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \end{pmatrix}$$

$$\Rightarrow \mathbf{M_1} = \mathbf{P\,A\,P^{-1}} = \mathbf{P\,A\,P^{T}} = \begin{pmatrix} x & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \end{pmatrix}$$

The Householder method

Suppose **A** is symmetric

$$\mathbf{A} = \begin{pmatrix} x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \end{pmatrix} \quad \Rightarrow \quad \mathbf{P\,A} = \begin{pmatrix} x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \end{pmatrix}$$

$$\Rightarrow \mathbf{M_1} = \mathbf{P\,A\,P^{-1}} = \mathbf{P\,A\,P^{T}} = \begin{pmatrix} x & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \end{pmatrix}$$

The Householder method

Next step, choose new Householder matrix of the form

$$
\mathbf{P} \;=\; \left(\begin{array}{c|cccccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & & & & & & \\
0 & & & & & & \\
0 & \multicolumn{6}{c}{\mathbf{I}_{N-1} - \mathbf{2}\; \mathbf{w}_{N-1}\mathbf{w}_{N-1}^{\mathbf{T}}} \\
0 & & & & & & \\
0 & & & & & & \\
0 & & & & & &
\end{array}\right)
\quad \text{to operate on } M_1 = \begin{pmatrix}
X & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & X & X & X & X & X & X \\
0 & X & X & X & X & X & X \\
0 & X & X & X & X & X & X \\
0 & X & X & X & X & X & X \\
0 & X & X & X & X & X & X \\
0 & X & X & X & X & X & X
\end{pmatrix}
$$

$$
\Rightarrow \mathbf{M_2 = P\, M_1\, P^{-1}} \;=\; \begin{pmatrix}
X & X & 0 & 0 & 0 & 0 & 0 \\
X & X & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & X & X & X & X & X \\
0 & 0 & X & X & X & X & X \\
0 & 0 & X & X & X & X & X \\
0 & 0 & X & X & X & X & X \\
0 & 0 & X & X & X & X & X
\end{pmatrix}
$$

The Householder method

After N Householder transformations, the matrix is tridiagonal

$$\mathbf{M_N} = \begin{pmatrix} \mathrm{X} & \mathrm{X} & 0 & 0 & 0 & 0 & 0 \\ \mathrm{X} & \mathrm{X} & \mathrm{X} & 0 & 0 & 0 & 0 \\ 0 & \mathrm{X} & \mathrm{X} & \mathrm{X} & 0 & 0 & 0 \\ 0 & 0 & \mathrm{X} & \mathrm{X} & \mathrm{X} & 0 & 0 \\ 0 & 0 & 0 & \mathrm{X} & \mathrm{X} & \mathrm{X} & 0 \\ 0 & 0 & 0 & 0 & \mathrm{X} & \mathrm{X} & \mathrm{X} \\ 0 & 0 & 0 & 0 & 0 & \mathrm{X} & \mathrm{X} \end{pmatrix} = \mathbf{P_N \ldots P_2 P_1\, A\, P_1}^{-1}\mathbf{P_2}^{-1} \ldots \mathbf{P_N}^{-1}$$

The Householder method: computational cost

You might think that each Householder transformation would require $O(N^3)$ operations (2 matrix multiplications), so the whole series would cost ~ $O(N^4)$ operations

However, we can rewrite the product $\mathbf{P\,A\,P}^{-1}$ as follows

$$\begin{aligned}
\mathbf{P\,A\,P}^{-1} = \mathbf{P\,A\,P} &= (\mathbf{I} - 2\,\mathbf{w\,w}^T)\,\mathbf{A}\,(\mathbf{I} - 2\,\mathbf{w\,w}^T) \\
&= (\mathbf{I} - 2\,\mathbf{w\,w}^T)\,(\mathbf{A} - 2\,\mathbf{A\,w\,w}^T) \\
&= (\mathbf{A} - 2\,\mathbf{w\,w}^T\mathbf{A} - 2\,\mathbf{A\,w\,w}^T + 4\,\mathbf{w\,w}^T\mathbf{Aw\,w}^T)
\end{aligned}$$

The Householder method: computational cost

Now define an N x 1 vector $\mathbf{v} = \mathbf{A}\,\mathbf{w}$

$$\mathbf{P}^{-1}\mathbf{A}\,\mathbf{P} = (\mathbf{A} - 2\mathbf{w}\mathbf{w}^T\mathbf{A} - 2\mathbf{A}\,\mathbf{w}\,\mathbf{w}^T + 4\,\mathbf{w}\,\mathbf{w}^T\,\mathbf{A}\,\mathbf{w}\,\mathbf{w}^T)$$
$$= (\mathbf{A} - 2\,\mathbf{w}\,\mathbf{v}^T - 2\,\mathbf{v}\,\mathbf{w}^T + 4\,\mathbf{w}\,\underbrace{\mathbf{w}^T\mathbf{v}}_{\text{scalar}}\,\mathbf{w}^T)$$

This calculation involves an N^2 operation to compute **v**, and a series of N^2 operations to compute, subtract and add various matrices

The reduction of the matrix to triadiagonal form therefore costs only $O(N^3)$ operations, not $O(N^4)$

Diagonalization of a symmetric tridiagonal matrix (*Recipes*, §11.3)

- The diagonalization of a symmetric tridiagonal matrix can be accomplished by a series of ~N 2 x 2 rotations
- Consider the rotation

(rotation through angle $\theta = \cos^{-1} C = \sin^{-1} S$ in the x_1 - x_2 plane)

$$R = \begin{pmatrix} C & S & 0 & 0 & 0 & 0 & 0 \\ -S & C & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Diagonalization of a symmetric tridiagonal matrix

- Applying this similarity transform to the triadiagonal matrix, we obtain

$$
R^T \begin{pmatrix} t & u & 0 & 0 & 0 & 0 & 0 \\ u & v & X & 0 & 0 & 0 & 0 \\ 0 & X & X & X & 0 & 0 & 0 \\ 0 & 0 & X & X & X & 0 & 0 \\ 0 & 0 & 0 & X & X & X & 0 \\ 0 & 0 & 0 & 0 & X & X & X \\ 0 & 0 & 0 & 0 & 0 & X & X \end{pmatrix} R = \begin{pmatrix} X & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & X & X & 0 & 0 & 0 & 0 \\ 0 & X & X & X & 0 & 0 & 0 \\ 0 & 0 & X & X & X & 0 & 0 \\ 0 & 0 & 0 & X & X & X & 0 \\ 0 & 0 & 0 & 0 & X & X & X \\ 0 & 0 & 0 & 0 & 0 & X & X \end{pmatrix}
$$

Diagonalization of a symmetric tridiagonal matrix

- To make the off-diagonal elements $(\mathbf{R}^{\mathbf{T}}\ \mathbf{M}_{\mathbf{N}}\ \mathbf{R})_{12} = (\mathbf{R}^{\mathbf{T}}\ \mathbf{M}_{\mathbf{N}}\ \mathbf{R})_{21}$ zero, we require

$$tSC + u(C^2 - S^2) - vSC = 0$$

$$\Rightarrow \tan\theta = \frac{(t - v) \pm \sqrt{(t - v)^2 + 4u^2}}{2u}$$