Holger Broman, August 19, 1999

Matrices, geometry, and mappings

Eigenvalues, eigenvectors and diagonalization

For a square matrix A, of dimension $N \times N$, the equation

$$\mathbf{y} = A\mathbf{x}$$

denotes a mapping from $\mathbb{R}^N$ to $\mathbb{R}^N$. In general, the vectors $\mathbf{x}$ and $\mathbf{y}$ are of course not collinear. The interesting question arises: do there exist vectors $\mathbf{g}$ such that, when mapped by A, they keep the original direction (or possibly reverse the direction)? Formulated differently, has the equation

$$A\mathbf{g} = \lambda\mathbf{g}$$

any solution for λ scalar and $\mathbf{g} \in \mathbb{R}^N$? The answer is yes, and that both λ and the elements of $\mathbf{g}$ in general are complex-valued, even if the elements of A are real-valued. Let us investigate the solutions by re-writing the equation:

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

Thus, $\mathbf{g}$ must be in the null-space of $A - \lambda I$. It also follows that the determinant of $A - \lambda I$ must be zero. As the determinant is an Nth degree polynomial in λ, it follows that A has exactly N eigenvalues. For eigenvectors, things are more complicated. If $\mathbf{g}$ is an eigenvector, then so is $c\mathbf{g}$, c scalar. Now, if we disregard this effect, then can we say that A has exactly N eigenvectors? The answer is no; there exist matrices that have less than N eigenvectors, but none having more than N.

———— o O o ————

We will proceed by demonstrating how certain matrices can be diagonalized. We will consider matrices such that their eigenvectors form a basis of $\mathbb{R}^N$. Thus

$$A\mathbf{g}_i = \lambda_i\mathbf{g}_i, \qquad i = 1, \ldots, N$$

and

$$\sum_{i=1}^{N} c_i\mathbf{g}_i = 0, \text{ only for } c_i = 0 \;\forall\; i.$$

It follows that

$$A\left[\mathbf{g}_1 \cdots \mathbf{g}_N\right] = \left[\mathbf{g}_1 \cdots \mathbf{g}_N\right] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{bmatrix}.$$

The proof is obtained by noting

- $A\,[\mathbf{g}_1 \cdots \mathbf{g}_N] = [A\mathbf{g}_1 \cdots A\mathbf{g}_N]$
- $[\mathbf{g}_1 \cdots \mathbf{g}_N] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{bmatrix} = [\lambda_1\mathbf{g}_1 \cdots \lambda_N\mathbf{g}_N]$

Introduce the notations

$$G = [\mathbf{g}_1 \cdots \mathbf{g}_N]$$

and

$$\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{bmatrix}.$$

Now, G is invertible as its columns are linearly independent. Thus

$$AG = G\Lambda \iff G^{-1}AG = \Lambda.$$

Finally, here is a theorem that gives a sufficient condition for a matrix to have a basis of eigenvectors:

> If the N eigenvalues of A are distinct, then A has a basis of eigenvectors.

The proof is by contradiction. Assume therefore that only M $(< N)$ eigenvectors are linearly independent. Use a permutation to assure that the M first are linearly independent. Then

$$\mathbf{g}_N = \sum_{m=1}^{M} c_m\,\mathbf{g}_m, \quad \text{not all } c_m \text{ zero}$$

As a consequence,

$$A\,\mathbf{g}_N = \sum_{m=1}^{M} c_m\,A\,\mathbf{g}_m,$$

which implies

$$\lambda_N\,\mathbf{g}_N = \sum_{m=1}^{M} c_m\,\lambda_m\,\mathbf{g}_m.$$

In addition,

$$\lambda_N\,\mathbf{g}_N = \sum_{m=1}^{M} c_m\,\lambda_N\,\mathbf{g}_m.$$

Subtract these two equations:

$$0 = \sum_{m=1}^{M} c_m(\lambda_m - \lambda_N)\mathbf{g}_m.$$

As $\lambda_m \neq \lambda_N \ \forall \ m$, and the set $\{\mathbf{g}_m\}_1^M$ is linearly independent, we have reached the desired contradiction as the only way we can satisfy the equation above is $c_m = 0 \ \forall \ m$.

Finally, a note on the eigenvalues of a matrix, and powers of that matrix. If λ is an eigenvalue of A, i.e.

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

then λ^K is an eigenvalue of A^K, as

$$A^K\mathbf{g} = A^{K-1}A\mathbf{g} = A^{K-1}\lambda\mathbf{g} = \lambda A^{K-1}\mathbf{g} = \ldots = \lambda^K\mathbf{g}.$$

Conversely, we can only make the obviously weaker statement. If λ is an eigenvalue of A^K, then an eigenvalue of A is to be found in the set $\left\{\lambda^{1/K}\right\}$.

——————— o O o ———————

Let us study symmetric matrices from an eigenvalue and eigenvector point of view. Let us also free ourselves from the requirement of real-valued entries in the matrix. The Hermitian transpose of a matrix is the transpose and complex conjugate of the matrix

$$A^H = \left(A^T\right)^*.$$

By definition, a Hermitian matrix fulfills

$$A = A^H.$$

A real-valued symmetric matrix is Hermitian.

For vectors with complex-valued entries, we extend the definitions of scalar product and norm:

> The scalar product between $\mathbf{x}$ and $\mathbf{y}$, both in $\mathbb{C}^N$, is $\mathbf{x}^H\mathbf{y} = \sum_{n=1}^{N} x_n^* \, y_n$

> The length (norm) of a complex vector is $\| \mathbf{x} \| = \left(\mathbf{x}^H \mathbf{x}\right)^{1/2}$

Here follow two important propositions:

> Hermitian matrices have real-valued eigenvalues.

Proof: Let

$$A\,\mathbf{g} = \lambda\,\mathbf{g}, \quad \mathbf{g} \neq 0.$$

Then

$$\mathbf{g}^H\;A\,\mathbf{g} = \lambda\,\mathbf{g}^H\,\mathbf{g} = \lambda\;\parallel \mathbf{g} \parallel^2$$

and

$$\left(\mathbf{g}^H\;A\,\mathbf{g}\right)^H = \mathbf{g}^H\;A^H\,\mathbf{g} = \mathbf{g}^H\,A\,\mathbf{g},$$

which shows that the left hand side, $\mathbf{g}^H\,A\,\mathbf{g} = \lambda \parallel \mathbf{g} \parallel^2$, is real-valued, as it is a scalar that equals its complex conjugate.

> Eigenvectors of Hermitian matrices that belong to different eigenvalues are orthogonal.

Proof: $A\,\mathbf{g}_1 = \lambda_1\,\mathbf{g}_1, \quad A\,\mathbf{g}_2 = \lambda_2\,\mathbf{g}_2$

$$\lambda_1\;\mathbf{g}_1^H\mathbf{g}_2 = (\lambda_1\;\mathbf{g}_1)^H\;\mathbf{g}_2 = (A\,\mathbf{g}_1)^H\;\mathbf{g}_2 =$$
$$= \mathbf{g}_1^H\;A^H\;\mathbf{g}_2 = \mathbf{g}_1^H\;A\,\mathbf{g}_2 = \mathbf{g}_1^H\;\lambda_2\;\mathbf{g}_2 = \lambda_2\;\mathbf{g}_1^H\;\mathbf{g}_2,$$

so we have

$$\lambda_1\left(\mathbf{g}_1^H\,\mathbf{g}_2\right) = \lambda_2\left(\mathbf{g}_1^H\,\mathbf{g}_2\right), \quad \lambda_1 \neq \lambda_2$$

$$\Longrightarrow$$

$$\mathbf{g}_1^H\,\mathbf{g}_2 = 0$$

Note: It is true that all Hermitian matrices have a full set of orthogonal eigenvectors, even for eigenvalues that have multiple occurrence. The proof is given in the appendix.

Note: It may be suggested that you have a look at the notes on "Norm-preserving linear mappings from $\mathbb{R}^N$ to $\mathbb{R}^N$" before you proceed here.

————— o O o —————

Now, for Hermitian matrices we have

$$G^{-1}\,A\,G = \Lambda = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_N)$$

where the columns of G constitute an orthonormal basis. For real-valued matrices A, we thus find

$$G^T\,G = \begin{pmatrix} \mathbf{g}_1^T \\ \vdots \\ \mathbf{g}_N^T \end{pmatrix} (\mathbf{g}_1 \ldots \mathbf{g}_N) = I$$

which greatly simplifies matters as G^{-1} is simple to compute: $G^{-1} = G^T$

For complex-valued matrices we find

$$G^H G = \ldots = I,$$

so that the diagonalization in both cases can be written

$$G^H A G = \Lambda.$$

Eigenvalues and eigenvectors are useful. One use is in the calculation of the principal axes and moments of inertia of a rigid body. Another is when solving differential equations by transform methods.

Please run the m-files eig1 and eig2 (after the Appendix).

Appendix: The spectral theorem[1]

Theorem: Let A be $N \times N$, Hermitian. Then, there exists a unitary matrix U (see norm-preserving linear mappings) and a real-valued diagonal matrix Λ so that

$$U^{-1}\,A\,U = \Lambda.$$

The columns of U are eigenvectors of A, the diagonal of Λ holds the eigenvalues of A.

The proof is by 'downwards' induction, and uses the following lemma.

Lemma: Let A be Hermitian and $\mathbf{g}$ an eigenvector, i.e. $A\,\mathbf{g} = \lambda\,\mathbf{g}$ for some λ. Then, if $\mathbf{v}$ is orthogonal to $\mathbf{g}$, then $A\,\mathbf{v}$ is orthogonal to $\mathbf{g}$.

Proof of lemma:

$$(A\,\mathbf{v})^H\mathbf{g} = \mathbf{v}^H\,A^H\,\mathbf{g} = \mathbf{v}^H\,A\,\mathbf{g} = \mathbf{v}^H\,\lambda\,\mathbf{g} = 0$$

Proof of the spectral theorem:

Any matrix has at least one eigenvector, as

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

has at least one solution. This follows from the fact that $\det(A - \lambda\,I) = 0$ is an Nth degree polynomial in λ.

Let $\mathbf{g}$ be a normalized eigenvector and construct a unitary matrix U_1 by letting $\mathbf{g}$ be its first column, and expand the rest of the columns any way you like (there is always an ON-base in $\mathbb{C}^N$). Then

$$\begin{aligned} U_1^{-1}\,A\,U_1 =& U_1^H\,A\,U_1 = \\ &= \begin{bmatrix} \mathbf{g}^H \\ \text{rest} \end{bmatrix} A\,[\mathbf{g} \quad \text{rest}] = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{bmatrix}. \end{aligned}$$

Of course, A_2 is Hermitian, $N-1 \times N-1$:

[1] Adapted from Jöran Bergh

$$\begin{bmatrix} \lambda_1^H & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_2^H & \\ 0 & & & \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{bmatrix} = \left(U_1^H\, A\, U_1\right)^H = U_1^H\, A^H\, U_1 =$$

$$= U_1^H\, A\; U_1 = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{bmatrix}$$

Now, apply the same argument to A_2:

$$U_2^{-1}\, A_2\, U_2 = \begin{bmatrix} \lambda_2 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_3 & \\ 0 & & & \end{bmatrix}$$

Combine

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & U_2^{-1} & \\ 0 & & & \end{bmatrix} U_1^{-1}\; A\; U_1 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & U_2 & \\ 0 & & & \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots\cdots & 0 \\ 0 & \lambda_2 & 0\cdots & 0 \\ \vdots & 0 & & \\ \vdots & \vdots & A_3 & \\ 0 & 0 & & \end{bmatrix}$$

Repeat in total N times, and note that a product of unitary matrices is unitary:

$$U^{-1}\, A\; U = \Lambda$$

The proof is completed.

```
% eig1.m
% eigenvalues and vectors
clear

% A random matrix
sprintf('A random matrix')
A=randn(2,2);
[V,D] = eig(A); D, V
pause

% symmetric implies real-valued eigenvalues
sprintf('A symmetric matrix')
B=A+A';
[V,D] = eig(B); D, V
pause

% 'square' implies non-negative definite
sprintf('A non-negative definite matrix')
C=A*A';
[V,D] = eig(C); D, V
pause

% eigenvalues to A*A
sprintf('Error in eigenvalues of squared matrix minus squared eigenvalues')
norm(sort(eig(A).*eig(A))-sort(eig(A*A)),'fro') % sort seems to work here
pause

% A pathological matrix
sprintf('A matrix that does not have a full set of eigenvectors')
[V,D]=eig([3 1;0 3]); D, V
```

```
% eig2.m

% Symmetric matrices, eigenvectors and ON-bases.
clear

% Generate a random, square matrix
N=100;
A=randn(N,N);

% Make it symmetric
A=A+A';

[V,D]=eig(A);

sprintf('The eigenvectors are orthonormal')
norm(V*V'-eye(N),'fro')
pause

sprintf('The inverse of the eigenvector matrix equals the transpose')
norm(V'-inv(V),'fro')
pause

sprintf('The matrix is diagonalized by the eigenvector matrix')
norm(D-V'*A*V,'fro')
pause

sprintf('The eigenvalues are real-valued')
norm(real(D)-D,'fro')
```