Math 52 0 - Linear algebra, Spring Semester 2012-2013
Dan Abramovich

The diagonalization of symmetric matrices.

This is the story of the eigenvectors and eigenvalues of a symmetric matrix A, meaning $A = A^T$.

It is a beautiful story which carries the beautiful name *the spectral theorem*:

Theorem 1 (The spectral theorem)**.** *If A is an $n \times n$ symmetric matrix then*

(1) All eigenvalues of A are real.
(2) A is orthogonally diagonalizable: $A = PDP^T$ where P is an orthogonal matrix and D is real diagonal.

There are immediate important consequences:

Corollary 2. *If A is an $n \times n$ symmetric matrix then*

(1) A has an orthogonal basis of eigenvectors $\mathbf{u}_i$.
(2) (spectral decomposition)

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T.$$

(3) The dimension of the λ eigenspace is the multiplicity of λ as a root of $\det(A - \lambda I)$.
(4) Different eigenspaces are orthgonal to each other

In fact a matrix A is orthogonally diagonalizable if and only if it is symmetric.

(The name *the spectral theorem* is inspired by another story of the inter-relationship of math and physics.)

The first part is directly proved:

Proposition 3. *The eigenvalues of a symmetric matrix are real.*

Proof. Suppose $A\mathbf{v} = \lambda\mathbf{v}$. We dot this with $\bar{\mathbf{v}}$, the complex conjugate:

$$\bar{\mathbf{v}} \cdot A\mathbf{v} = \bar{\mathbf{v}} \cdot \lambda\mathbf{v}$$

The right hand side is $\lambda(|v_1|^2 + \cdots + |v_n|^2)$, where v_i are the complex entries of $\mathbf{v}$. Then λ is real if and only if the right hand side is real, if and only if $\bar{\mathbf{v}} \cdot A\mathbf{v} = \overline{\bar{\mathbf{v}} \cdot A\mathbf{v}}$. Now

$$\overline{\bar{\mathbf{v}} \cdot A\mathbf{v}} = \overline{\bar{\mathbf{v}}^T A\mathbf{v}} = \overline{\mathbf{v}^T \cdot A^T \bar{\mathbf{v}}} = \bar{\mathbf{v}} \cdot A^T\mathbf{v} = \bar{\mathbf{v}} \cdot A\mathbf{v}$$

which is what we needed. ♣

Here are key geometric facts:

Proposition 4. *If A is symmetric and $\mathcal{B} = \{\mathbf{u}_1 \ldots, \mathbf{u}_n\}$ is an orthonormal basis then $[A]_{\mathcal{B}}$ is symmetric.*

Proof. We know that

$$[A]_{\mathcal{B}} = P_{\mathcal{B}}^{-1} A P_{\mathcal{B}} = P_{\mathcal{B}}^T A P_{\mathcal{B}}.$$

Its transpose is

$$([A]_{\mathcal{B}})^T = P_{\mathcal{B}}^T A^T (P_{\mathcal{B}}^T)^T = P_{\mathcal{B}}^T A P_{\mathcal{B}} = [A]_{\mathcal{B}}.$$

♣

Proposition 5. *If P and Q are orthogonal matrices then PQ is also orthogonal.*

Proof.

$$(PQ)^T(PQ) = (Q^T P^T)(PQ) = Q^T(P^T P)Q = Q^T Q = I.$$

♣

Proof of the spectral theorem (Theorem 1). We argue by induction. The case $n = 1$ is trivial. Assume $n > 1$ and the result known for $n - 1$.

Suppose λ_1 is an eigenvalue and u_1 is a corresponding eigenvector. By Proposition 3 the eigenvalue λ is real. Normalizing, we may assume $\|\mathbf{u}_1\| = 1$.

By the orthogonalization procedure (Gram-Schmidt) we may complete this to an orthonormal basis $\mathcal{B} = \{\mathbf{u}_1 \ldots, \mathbf{u}_n\}$. Then $[A]_\mathcal{B}$ is symmetric. Since $A\mathbf{u}_1 = \lambda_1\mathbf{u}_1$ we get

$$[A]_\mathcal{B} = \left[\begin{array}{c|ccc} \lambda_1 & * & * & * \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right].$$

By Proposition 4 this is symmetric, so A_1 is symmetric and

$$[A]_\mathcal{B} = \left[\begin{array}{c|ccc} \lambda_1 & 0 & 0 & 0 \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right].$$

By induction $A_1 = P_1 D_1 {P_1}^T$ withP_1 orthogonal and D_1 real diagonal. Writing

$$Q_1 = \left[\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ \vdots & & P_1 & \\ 0 & & & \end{array}\right] \qquad \text{and} \qquad D = \left[\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ \vdots & & D_1 & \\ 0 & & & \end{array}\right]$$

we have $[A]_\mathcal{B} = Q_1 D Q_1^T$, and Q_1 is orthogonal

So $A = P_\mathcal{B}[A]_\mathcal{B}P_\mathcal{B}^T = P_\mathcal{B}Q_1DQ_1^TP_\mathcal{B}^T = P_\mathcal{B}Q_1D(P_\mathcal{B}Q_1)^T$. By Proposition 5 the matrix $P = P_\mathcal{B}Q_1$ is orthogonal, so $A = PDP^T$ is as required. ♣

Proof of Corollary 2. (1) The columns of P are a basis of eigenvectors, which is orthonormal since P is an orthogonal matrix.

(2) This follows by writing out $P = [\mathbf{u}_1 \cdots \mathbf{u}_n]$. Note $\mathbf{u}_i\mathbf{u}_i^T = \mathrm{proj}_{\mathbf{u}_i}$, so this makes sense geometrically.

(3) This follows since it is true for D.

(4) If $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ and $A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$ then

$$(A\mathbf{v}_1) \cdot \mathbf{v}_2 = \mathbf{v}_2^T A^T \mathbf{v}_1 = (A\mathbf{v}_1)^T \mathbf{v}_2 = \mathbf{v}_1^T A\mathbf{v}_2 = \mathbf{v}_1 \cdot (A\mathbf{v}_2)$$

so $\lambda_1\mathbf{v}_1 \cdot \mathbf{v}_2 = \lambda_2\mathbf{v}_2 \cdot \mathbf{v}_2$.

Since $\lambda_1 \neq \lambda_2$ we get $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.

The last statement: assume $A = PDP^T$ with P orthogonal. Then

$$A^T = (P^T)^T D^T P^T = PDP^T = A.$$

♣

The singular value decomposition

geometric discussion:

If A is symmetric then, in some orthogonal basis we get $[A]_{\mathcal{B}}$ diagonal.

Consider the n-sphere

$$\mathbb{S}^n = \{\mathbf{x} : \|\mathbf{x}\| = 1\}.$$

Then it is clear what A does to the sphere: it makes it into an ellipsoid.

←draw

Possibly crushed ellipsoid, if A is not invertible, but let's assume it is.

Written in diagonal form, if

$$\mathbf{x} = c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n$$

is on the sphere then

$$c_1^2 + \cdots + c_n^2 = 1,$$

and

$$\begin{aligned} A\mathbf{x} &= \lambda_1 c_1 \mathbf{u}_1 + \cdots + \lambda_n c_n \mathbf{u}_n \\ &= y_1\mathbf{u}_1 + \cdots + y_n\mathbf{u}_n. \end{aligned}$$

So the image is defined by

$$\frac{y_1^2}{\lambda_1^2} + \cdots + \frac{y_n^2}{\lambda_n^2} = 1.$$

The directions of $\mathbf{u}_i$ are the *axis directions* or *principal directions* of the ellipsoid, and $|\lambda_i|$ are the lengths of the *semi-axes*. So the axes, or diameters, are of lengths $2|\lambda_i|$.

It is also easy to see that for any A, even rectangular, the image $A(\mathbb{S}^n)$ is an ellipsoid. Try

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

It is not as easy to see what the axes are!

←do it, draw

The most important are the maximal and minimal semi-axes, since they say how much a vector can be stretched or shrunk.

Key observation:

$$\|A\mathbf{v}\|^2 = \mathbf{v}^T(A^TA)\mathbf{v},$$

and the matrix A^TA is symmetric!

If $\mathbf{v}_i$ are unit eigenvectors of A^TA with eigenvalue λ_i, then

$$\|A\mathbf{v}_i\|^2 = \lambda_i\|\mathbf{v}_i\|^2 = \lambda_i,$$

so $\lambda_i \geq 0$ and

$$\|A\mathbf{v}_i\| = \sqrt{\lambda_i}.$$

The numbers $\sigma_i = \sqrt{\lambda_i}$ are called the *singular values* of A.

They are the lengths of $A\mathbf{v}_i$

do the example→ These give axes of the ellipsoid!

Theorem 6. *Suppose*

$$\lambda_1 \dots \lambda_r \neq 0$$

and the others (if any)

$$\lambda_{r+1} = \cdots = 0.$$

Then

$$A\mathbf{v}_1, \dots, A\mathbf{v}_r$$

is an orthogonal basis for **Col** A*.*

Write $\Sigma = \left[\begin{array}{ccc|c} \sigma_1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & \sigma_r & 0 \\ \hline 0 & \cdots & 0 & 0 \end{array}\right]$

Theorem 7 (Singular value decomposition)**.** *There is an $m \times m$ orthogonal U, and an $n \times n$ orthogonal V, such that*

geometric interpretation→

$$A = U\Sigma V^T.$$

algorithmic→

from before→

signular values→

Proof. The matrix V is simply $[\mathbf{v}_1 \ldots \mathbf{v}_n]$.

The matrix Σ is as indicated.

For the matrix U take normalizations

$$\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$$

for $i = 1, \ldots, r$ and complete to an orthonormal basis.

Then check that $AV = U\Sigma$. ♣

do the example→

Interpretations:

(1) $A = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$

(2) the matrix of A in the bases $\{\mathbf{v}_i\}$ for source $\mathbb{R}^n$ and $\mathbf{u}_i$ for target $\mathbb{R}^m$ is the diagonal matrix Σ.

(3) A is a rotation followed by diagonal stretching followed by another rotation.