Symmetric Matrices & the Spectral Theorem

Recall: S is symmetric if $S=S^T$

(in particular, it is square)

Primary Example: the matrix of column dot products

$$\boxed{S=A^TA \quad \text{for any matrix } A}$$

Diagonalization works out very nicely for symmetric matrices; this is summed up in the spectral theorem, our main topic for today. The spectral theorem is the main ingredient in the SVD also.

Eg: $S=\frac{1}{9}\begin{pmatrix} 5 & -8 & 10 \\ -8 & 11 & 2 \\ 10 & 2 & 2 \end{pmatrix}$

[DEMO]

What do you notice about the eigenspaces?

Fact ①: Let S be an $n\times n$ symmetric matrix & $u,v\in\mathbb{R}^n$.

Then $\boxed{v\cdot(Sw) = (Sv)\cdot w}$.

Proof: $v\cdot(Sw) = v^T(Sw) = (v^TS)w = (S^Tv)^Tw \overset{S=S^T}{=} (Sv)^Tw$

$= (Sv)\cdot w$ //

Fact ②: Eigenvectors of S with different eigenvalues are orthogonal.

Proof: Say $Sv_1 = \lambda_1 v_1$ $\quad Sv_2 = \lambda_2 v_2$ $\quad \lambda_1 \neq \lambda_2$. Then

$$v_1\cdot(Sv_2) = v_1\cdot(\lambda_2 v_2) = \lambda_2(v_1\cdot v_2)$$

$$① \parallel$$

$$(Sv_1)\cdot v_2 = (\lambda_1 v_1)\cdot v_2 = \lambda_1(v_1\cdot v_2)$$

So $\lambda_1(v_1\cdot v_2) = \lambda_2(v_1\cdot v_2) \implies (\lambda_1-\lambda_2)(v_1\cdot v_2) = 0$

$\overset{\lambda_1\neq\lambda_2}{\implies} v_1\cdot v_2 = 0$, which means $v_1 \perp v_2$. //

Fact ③: All eigenvalues of S are real numbers.

Proof: Say $Sv = \lambda v$, $v \neq 0$, and λ is not real, so $\lambda \neq \bar\lambda$ ($a+ib \neq a-ib$ when $b \neq 0$). We know $S\bar v = \bar\lambda \bar v$, so ② $\implies v\cdot\bar v = 0$ (they are eigenvectors with different eigenvalues). But

$$v = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \implies \bar v = \begin{pmatrix} \bar z_1 \\ \vdots \\ \bar z_n \end{pmatrix}$$

$$\implies v\cdot\bar v = z_1\bar z_1 + z_2\bar z_2 + \cdots + z_n\bar z_n$$

$$= |z_1|^2 + |z_2|^2 + \cdots + |z_n|^2,$$

which is > 0 because $v \neq 0$. So no such non-real eigenvalue can exist. //

Fact ④: If S is symmetric and λ is an eigenvalue of S, then $AM(\lambda) = GM(\lambda)$.

(I can't prove this without ideas from abstract vector spaces.)

Consequence of ②③④:

Any symmetric matrix S is diagonalizable over $\mathbb{R}$. Moreover, it has an orthonormal eigenbasis.

Eg: $S = \frac{1}{9}\begin{pmatrix} 5 & -8 & 10 \\ -8 & 11 & 2 \\ 10 & 2 & 2 \end{pmatrix}$ $\quad p(\lambda) = -(\lambda-1)(\lambda+1)(\lambda-2)$

Compute eigenvectors:

$$\lambda = 1 \qquad \lambda = -1 \qquad \lambda = 2$$

$$w_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \qquad w_2 = \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix} \qquad w_3 = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$$

Check orthogonality:

$$w_1 \cdot w_2 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix} = 0 \quad w_1 \cdot w_3 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} = 0 \quad w_2 \cdot w_3 = \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} = 0$$

✓

Orthonormal Eigenbasis:

This shows $\{w_1, w_2, w_3\}$ is an orthogonal eigenbasis. To make it orthonormal, we divide by the lengths:

$$\|w_1\| = \sqrt{9} = 3 \qquad \|w_2\| = 3 \qquad \|w_3\| = 3$$

$$\Longrightarrow \left\{ \frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \frac{1}{3}\begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}, \frac{1}{3}\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \right\}$$

is an orthonormal eigenbasis.

Matrix form:

orthogonal matrix ↓

$$S = QDQ^{-1} = QDQ^{T} \quad \text{for } Q = \frac{1}{3}\begin{pmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{pmatrix} \; D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Recall: A square matrix Q with orthonormal columns is called orthogonal. This means

$$QQ^T = I_n \quad \text{so} \quad Q^T = Q^{-1}.$$

To summarize:

Spectral Theorem: A real symmetric matrix S has an orthonormal basis of real eigenvectors:

$$S = QDQ^T$$

for Q orthogonal and D diagonal.

Fast-Forward: Apply the spectral theorem to $S = A^TA$ and you get 90% of the SVD for A.

NB: Conversely, if $S = QDQ^T$ for D diagonal then S must be symmetric, because

$$S^T = (QDQ^T)^T = Q^{TT}D^TQ^T = QDQ^T = S.$$

In other words, if S has an orthonormal eigenbasis then S is necessarily symmetric! More on HW.

Eg: $S = \begin{pmatrix} -1 & 1 & 2 \\ 1 & -1 & 2 \\ 2 & 2 & 2 \end{pmatrix}$ $\qquad p(\lambda) = -(\lambda-4)(\lambda+2)^2$

Eigenspace Bases:

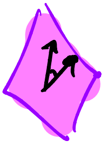
$\lambda = 4: \left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\}$ $\qquad \lambda = -2: \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \right\}$ ←

NB: Since $AM(-2) = 2$, we must get 2 vectors here.

Check orthogonality:

$$\begin{pmatrix}1\\1\\2\end{pmatrix}\cdot\begin{pmatrix}-1\\1\\0\end{pmatrix}=0 \qquad \begin{pmatrix}1\\1\\2\end{pmatrix}\cdot\begin{pmatrix}-2\\0\\1\end{pmatrix}=0 \qquad \begin{pmatrix}-1\\1\\0\end{pmatrix}\cdot\begin{pmatrix}-2\\0\\1\end{pmatrix}=2 \overset{???}{\neq 0}$$

Doesn't this contradict Fact ②?



No! These eigenvectors have the same eigenvalue $\lambda=-2$. In fact, since the (-2)-eigenspace is a plane, there will be plenty of non-orthogonal (-2)-eigenvectors!

[DEMO]

Ok, so how do we produce an orthonormal eigenbasis?

Gram-Schmidt to the rescue!

$$w_1=\begin{pmatrix}1\\1\\2\end{pmatrix} \qquad w_3=\begin{pmatrix}-2\\0\\1\end{pmatrix}-\frac{\begin{pmatrix}-2\\0\\1\end{pmatrix}\cdot\begin{pmatrix}-1\\1\\0\end{pmatrix}}{\begin{pmatrix}-1\\1\\0\end{pmatrix}\cdot\begin{pmatrix}-1\\1\\0\end{pmatrix}}\begin{pmatrix}-1\\1\\0\end{pmatrix}=\begin{pmatrix}-1\\-1\\1\end{pmatrix}$$

$$w_2=\begin{pmatrix}-1\\1\\0\end{pmatrix}$$

(We only had to do G-S on $\left\{\begin{pmatrix}-1\\1\\0\end{pmatrix},\begin{pmatrix}-2\\0\\1\end{pmatrix}\right\}$ because they're already $\perp w_1$.)

Check orthogonality (again):

$$\begin{pmatrix}1\\1\\2\end{pmatrix}\cdot\begin{pmatrix}-1\\1\\0\end{pmatrix}=0 \qquad \begin{pmatrix}1\\1\\2\end{pmatrix}\cdot\begin{pmatrix}-1\\-1\\1\end{pmatrix}=0 \qquad \begin{pmatrix}-1\\1\\0\end{pmatrix}\cdot\begin{pmatrix}-1\\-1\\1\end{pmatrix}=0 \quad ✓$$

So $S=QDQ^T$ for

$$Q=\begin{pmatrix}1/\sqrt{6} & -1/\sqrt{2} & -1/\sqrt{3}\\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3}\\ 2/\sqrt{6} & 0 & 1/\sqrt{3}\end{pmatrix} \qquad D=\begin{pmatrix}4&0&0\\0&-2&0\\0&0&-2\end{pmatrix}$$

$$\begin{matrix}\shortparallel & \shortparallel & \shortparallel\\ w_1/\|w_1\| & w_2/\|w_2\| & w_3/\|w_3\|\end{matrix}$$

How to Orthogonally Diagonalize a Real, Symmetric Matrix:

(1) Diagonalize S. (it is necessarily diagonalizable)

(2) Run Gram-Schmidt on eigenspace bases with multiplicity ≥2.

(3) Divide by lengths.

The result is an orthonormal eigenbasis.

Eg: $S = \frac{1}{2}\begin{pmatrix}5 & -1\\ -1 & 5\end{pmatrix}$ $\quad p(\lambda) = \lambda^2 - 5\lambda + 6 = (\lambda-2)(\lambda-3)$

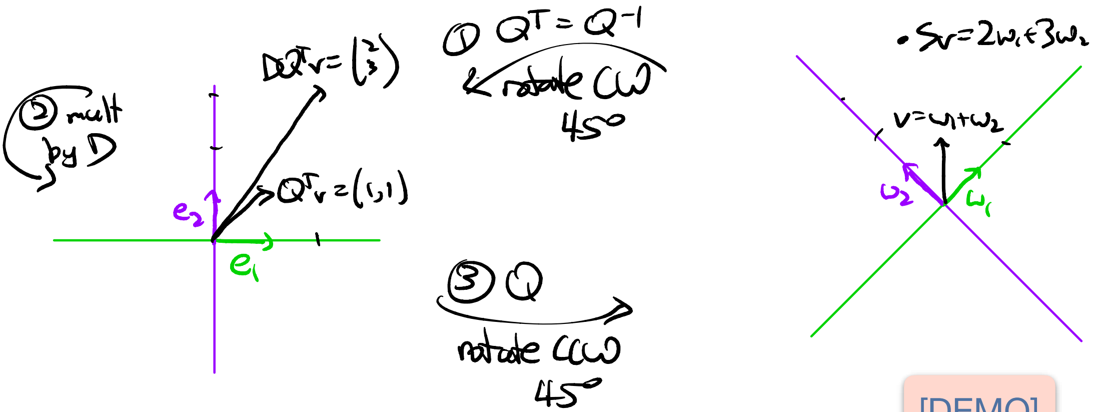
$\lambda_1 = 2 \quad w_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \qquad \lambda_2 = 3 \quad w_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix}$

So $S = QDQ^T$ for

$$Q = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix} \qquad D = \begin{pmatrix}2 & 0\\ 0 & 3\end{pmatrix}$$

Let's draw a picture of this.

$Qe_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \qquad Qe_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix}$ } Q rotates CCW by 45°



[DEMO]

This is the same picture as before, except that it's much easier to visualize multiplication by Q and $Q^{-1} = Q^T$.

Think: Q preserves lengths and angles.

Multiplication by an orthogonal matrix is basically a rotation or a reflection.

More on this when we do the SVD.

Exercise (outer product form):

If $\{u_1, u_2, \dots, u_n\}$ is an orthonormal eigenbasis of S and $Su_i = \lambda_i u_i$, so $S = QDQ^T$ for

$$Q = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix} \quad D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

then

$$S = \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \cdots + \lambda_n u_n u_n^T.$$

This is almost the SVD of A.

Compare: If P_V is the projection matrix onto a subspace V, then

$$P_V = u_1 u_1^T + u_2 u_2^T + \cdots + u_d u_d^T$$

if $\{u_1, u_2, \dots, u_d\}$ is an orthonormal basis of V.

In fact, this is a special case of the exercise above: V is the 1-eigenspace, and the other eigenvalue is 0. (Recall that P_V is symmetric.)

Positive-Definite Symmetric Matrices

Recall: $S = A^TA$ is a very important example of a symmetric matrix!

Observation: If v is an eigenvector of S with eigenvalue λ, then

$$v\cdot(Sv) = v\cdot(\lambda v) = \lambda(v\cdot v) = \lambda\|v\|^2$$

and

$$v\cdot(Sv) = v^TSv \overset{S=A^TA}{=} v^TA^TAv = (Av)^T(Av)$$

$$= (Av)\cdot(Av) = \|Av\|^2.$$

So:

$$\boxed{\lambda\|v\|^2 = \|Av\|^2}$$

Consequences:

- If λ is an eigenvalue of $S = A^TA$ then $\lambda \geq 0$.

If $\lambda = 0$ then $v \in \text{Nul}(A) = 0\text{-eigenspace}$, so:

- If A has full column rank then $\lambda > 0$.

Hence A^TA only has positive eigenvalues when A has FCR.

This condition is very important, so it has a name.

Def: A symmetric matrix S is called:

Example

- positive-definite (+ve-def)
  if its eigenvalues are all positive.

$$Q\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}Q^T$$

- positive-semidefinite (+ve-semidef)
  if its eigenvalues are all nonnegative.
  (this allows $\lambda = 0$ as well)

$$Q\begin{pmatrix}1&0&0\\0&2&0\\0&0&0\end{pmatrix}Q^T$$

- indefinite
  if it has positive and negative eigenvalues.

$$Q\begin{pmatrix}1&0&0\\0&-2&0\\0&0&0\end{pmatrix}Q^T$$

- negative-definite, etc.

NB: A positive-definite matrix is also positive-semidefinite because $\lambda > 0 \Rightarrow \lambda \geq 0$.

Amazingly, it is possible to check positive-definiteness without computing any eigenvalues, thanks to the following incredible theorem.

Criteria for Positive-Definiteness:

Let S be a symmetric matrix.

The Following Are Equivalent:

(1) S is positive-definite

(2) [Positive Energy Criterion]

$x^TSx > 0$ for all $x \neq 0$.

(3) [Determinant Criterion] The determinants of all n upper-left submatrices are positive:

$$S = \begin{pmatrix} 7 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 5 \end{pmatrix} \rightsquigarrow \det\begin{pmatrix} 7 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 5 \end{pmatrix} > 0$$

$$\det\begin{pmatrix} 7 & 2 \\ 2 & 6 \end{pmatrix} > 0$$

$$\det\begin{pmatrix} 7 \end{pmatrix} > 0$$

(4) [Cholesky Decomposition]

$S = A^TA$ for a matrix A with full column rank.

(5) [LDL^T Decomposition]

S has an LU decomposition where U has positive diagonal entries.

(no row swaps needed!)

(5) is fastest: it's just an elimination problem!

Remarks:

(2) In physics, $x^TSx = \langle x|T|x\rangle$ sometimes measures the energy of a system.

(2 ⇒ 1) If λ is an eigenvalue with eigenvector v then

$$0 < v^TSv = v\cdot(Sv) = v\cdot(\lambda v) = \lambda\|v\|^2$$

$$\Rightarrow \lambda > 0$$ ✓

(1 ⇒ 2) Let $x\in\mathbb{R}^n$, $x\neq 0$. We can orthogonally diagonalize: $S=QDQ^T$ where

$$D=\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix} \qquad \lambda_i>0$$

Since $\mathrm{Nul}(Q^T)=0$ we have $y=Q^Tx\neq 0$.

Then

$$x^TSx = x^TQDQ^Tx = (Q^Tx)^TD(Q^Tx)$$

$$= y^TDy = (y_1\cdots y_n)\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$$

$$= \lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2>0$$ ✓

(3) This is proved in the LDL^T supplement (L22).

(4 ⇒ 1) We did this above.

(1 ⇒ 4) We'll discuss the Cholesky Decomposition next time.

(5) We'll discuss the LDL^T Decomposition next time.

There are analogous

Criteria for Positive-Semidefiniteness:

Let S be a symmetric matrix.

The Following Are Equivalent:

(1) S is positive-semidefinite

(2) $x^TSx \geq 0$ for all $x \neq 0$

(3) The determinants of all n upper-left submatrices are nonnegative.

(4) $S = A^TA$ for a matrix A ~~with full column rank.~~

Recap: If A is any matrix, then $S = A^TA$ has nonnegative eigenvalues, so it is positive-semidefinite.