Symmetric Matrices & the Spectral Theorem

Recall: S is symmetric if $S = S^T$ ($\Rightarrow$ square)

Super-important example: the matrix of column dot products

$S = A^TA$ for any matrix A $\quad [(A^TA)^T = A^TA^{TT} = A^TA]$

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?

Observation 0: for any vectors v and w,

$$v\cdot(Sw) = v^TSw = (S^Tv)^Tw = (Sv)^Tw = (Sv)\cdot w$$

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

Observation 1:

Eigenvectors of S with different eigenvalues are orthogonal.

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

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

$$\parallel$$

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

$$\lambda_1 v_1 \cdot v_2 = \lambda_2 v_1 \cdot v_2 \Longrightarrow (\lambda_1 - \lambda_2) v_1 \cdot v_2 = 0$$
$$\xrightarrow{\lambda_1 \neq \lambda_2} v_1 \cdot v_2 = 0 \quad ✓$$

Observation 2:
All eigenvalues of S are real.

Proof: Say $Sv = \lambda v$ and λ is not real.
Then $\lambda \neq \bar{\lambda}$. Conjugate eigenvalue: $S\bar{v} = \bar{\lambda}\bar{v}$.
Observation 1 $\Longrightarrow v \cdot \bar{v} = 0$. But

$$v = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \qquad \bar{v} = \begin{pmatrix} \bar{z}_1 \\ \vdots \\ \bar{z}_n \end{pmatrix}$$

$$\Longrightarrow v \cdot \bar{v} = z_1 \bar{z}_1 + \cdots + z_n \bar{z}$$
$$= |z_1|^2 + \cdots + |z_n|^2 > 0$$

So this can't happen. ✓

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

(The proof requires ideas from abstract linear algebra)

Consequence: S is diagonalizable over the real numbers! Moreover, there is an orthonormal eigenbasis.

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

Eigenvectors:

$\lambda = 1 \rightsquigarrow w_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$

$\lambda = -1 \rightsquigarrow w_2 = \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}$

$\lambda = 2 \rightsquigarrow w_3 = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$

Check:

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

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

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

So $\{w_1, w_2, w_3\}$ is an orthogonal eigenbasis.
To make it orthonormal, you have to divide by the lengths to make them unit vectors:

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

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

is an orthonormal eigenbasis.

Matrix form:

orthogonal

$S = QDQ^{-1}$ for $Q = \frac{1}{3}\begin{pmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{pmatrix}$

$= QDQ^T \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

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

$$Q^TQ = I_n \implies Q^T = Q^{-1}.$$

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

$$S = QDQ^T$$

for an orthogonal matrix Q and a diagonal matrix D.

Fast-Forward: The SVD is basically the spectral theorem as applied to $S = A^TA$.

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

Eigenspaces:

$\lambda = 4 \rightsquigarrow \mathrm{Span}\left\{\begin{pmatrix}1\\1\\2\end{pmatrix}\right\}$

$\lambda = -2 \rightsquigarrow \mathrm{Span}\left\{\begin{pmatrix}-1\\1\\0\end{pmatrix}, \begin{pmatrix}-2\\0\\1\end{pmatrix}\right\}$

Check: $\begin{pmatrix}1\\1\\2\end{pmatrix}\cdot\begin{pmatrix}-1\\1\\0\end{pmatrix} = 0 \quad \begin{pmatrix}1\\1\\2\end{pmatrix}\cdot\begin{pmatrix}-2\\0\\1\end{pmatrix} = 0$

$\begin{pmatrix}-1\\1\\0\end{pmatrix}\cdot\begin{pmatrix}-2\\0\\1\end{pmatrix} = 2 \neq 0!$

That's ok — $\begin{pmatrix}-1\\1\\0\end{pmatrix}$ and $\begin{pmatrix}-2\\0\\1\end{pmatrix}$ have the same eigenvalue.

So how do we produce an orthonormal eigenbasis?
Have to use Gram-Schmidt to find an orthonormal basis of the -2-eigenspace.

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

$$w_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \qquad = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} - \frac{2}{2} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

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

So $\left\{ \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right\}$ is an orthonormal eigenbasis, and $S = QDQ^T$ for

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

Procedure to Orthogonally Diagonalize a Real Symmetric Matrix S:

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

(2) Normalize your eigenvectors / run Gram-Schmidt if $GM(\lambda) \geq 2$.

(3) Put them together $\rightsquigarrow$ orthonormal eigenbasis!

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

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

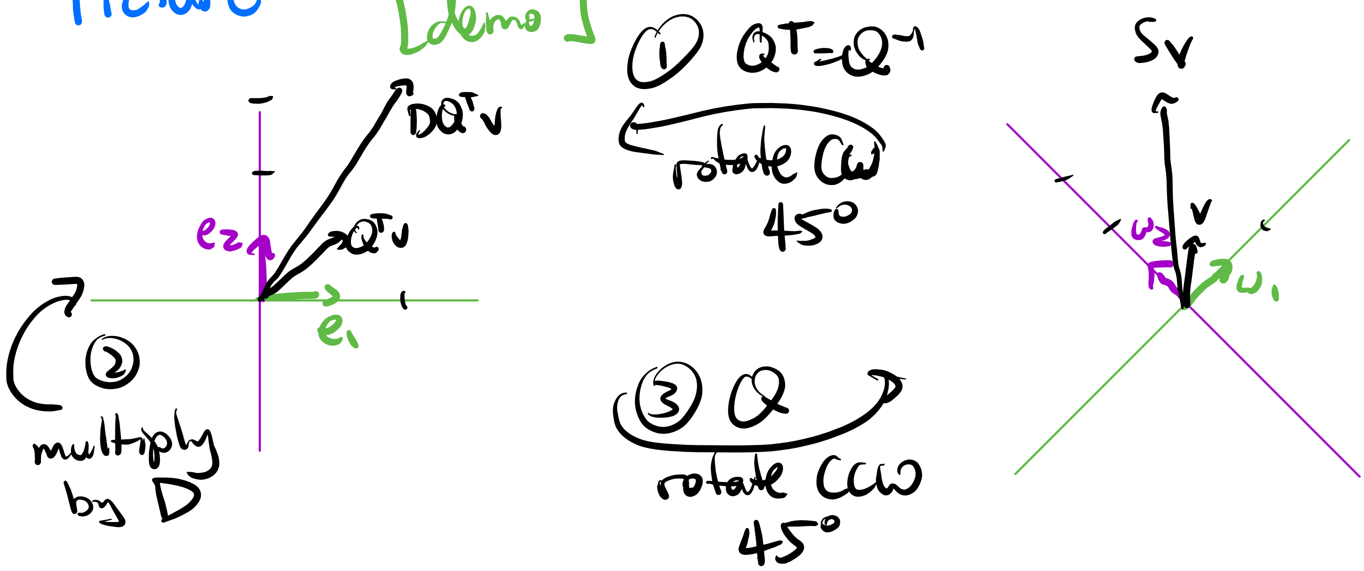
So $S = QDQ^T$ for $Q = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ $D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$

NB: $Q = \begin{pmatrix} \cos(45^\circ) & -\sin(45^\circ) \\ \sin(45^\circ) & \cos(45^\circ) \end{pmatrix}$

so Qx = (rotate x CCW 45°)

Picture:

[demo]



The picture is the same as before, but it's easier to visualize multiplying by the orthogonal matrix Q (it preserves lengths & angles).

Exercise (outer product form):

If $\{u_1, \dots, u_n\}$ 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

$$\boxed{S = \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \cdots + \lambda_n u_n u_n^T}$$

Compare: if P_V is a projection matrix, you can write $P_V = QQ^T$ for $Q = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_d \\ | & & | \end{pmatrix}$ $d = \dim(V)$.

$$\rightsquigarrow P_V = u_1 u_1^T + \cdots + u_d u_d^T.$$

(This is a special case: $\lambda_1 = \cdots = \lambda_d = 1$ and $\lambda_{d+1} = \cdots = \lambda_n = 0$. Recall P_V is symmetric!)

Positive-Definite Symmetric Matrices

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

Observation: If λ is an eigenvalue of $S = A^TA$ with eigenvector v then

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

$$v \cdot Sv = v^T S v = v^T A^T A v = (Av)^T (Av) = (Av) \cdot (Av) = \|Av\|^2$$

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

Consequence: If λ is an eigenvalue of $S = A^TA$ then $\lambda \geq 0$. Moreover, $\lambda = 0 \iff \|Av\| = 0 \iff v \in \text{Nul}(A)$, so if A has full column rank then $\lambda > 0$.

Thus A^TA has only positive eigenvalues when A has full column rank. This condition is so important that it has a name.

Def: A symmetric matrix S is called

- **positive-definite** if all its eigenvalues are positive.
- **positive-semidefinite** if all its eigenvalues are non-negative.
  (positive-semidefinite allows $\lambda=0$ as well.)
- **indefinite** if it has positive and negative eigenvalues.

NB: A positive-definite matrix is also positive-semidefinite!

$$\lambda>0 \implies \lambda\geq 0$$

Fast-forward: This will be important for solving quadratic optimization problems.

Eg:

- $Q\begin{pmatrix}3&0\\0&2\end{pmatrix}Q^T$ is positive-definite
- $Q\begin{pmatrix}3&0\\0&0\end{pmatrix}Q^T$ is positive-semidefinite
- $Q\begin{pmatrix}3&0\\0&-2\end{pmatrix}Q^T$ is indefinite.

Positive-definiteness is an important condition. We really want to be able to check it without computing eigenvalues.

Criteria for Positive-Definiteness:

Let S be a symmetric matrix.
The Following Are Equivalent:

(1) S is positive-definite

(2) $x^T S x > 0$ for all $x \neq 0$ ("positive energy")

(3) The determinants of all n upper-left submatrices are positive:

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

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

$$\det(7) > 0$$

(4) $S = A^T A$ for a matrix A with full column rank

(5) S has an LU decomposition where U has positive diagonal entries.
(no row swaps needed!)

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

Remarks:

(2) In physics, $x^T S x$ sometimes measures the energy of a system.

In any case, if v is an eigenvector with eigenvalue λ then

$$v^T S v = v \cdot \lambda v = \lambda \|v\|^2$$

so (2) $\Rightarrow$ $\lambda > 0$ for all λ, so (2) $\Rightarrow$ (1).

Conversely, (1) $\Rightarrow$ (2) because if $x \neq 0$ then $Q^T x \neq 0$, so if $Q^T x = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ then

$$x^T S x = x^T Q D Q^T x = (Q^T x)^T D (Q^T x)$$

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

$$= \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 > 0.$$

(3) Determinants are magic.

(Also see the LDL^T supplement.)

(4) (4) $\Rightarrow$ (1): we did this above.

(1) $\Rightarrow$ (4): This is the Cholesky decomposition: next time

(5) This is the LDL^T decomposition: next time.

Criteria for Positive-Semidefiniteness:

Let S be a symmetric matrix.
The following are equivalent:

(1) S is positive-semidefinite

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

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

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

Consequence: If A is any matrix then $A^T A$ is positive-semidefinite. In particular, it has nonnegative eigenvalues.