Application: Constrained Optimization (cont'd)

Recall: a quadratic form is
$$q(x_1,\dots,x_n)=\sum a_{ij}x_ix_j \iff q(x)=x^TSx,\ S \text{ symmetric}$$

Eg: $q(x_1,x_2,x_3)=a_{11}x_1^2+a_{22}x_2^2+a_{33}x_3^2+a_{12}x_1x_2+a_{13}x_1x_3+a_{23}x_2x_3$

then $q(x)=x^TSx$ for $S=\frac{1}{2}\begin{pmatrix}2a_{11} & a_{12} & a_{13}\\ a_{12} & 2a_{22} & a_{23}\\ a_{13} & a_{23} & 2a_{33}\end{pmatrix}$

Quadratic Optimization Problem:

Find min/max values of $q(x)$ subject to $\|x\|=1$.

Procedure for Solving a Quadratic Optimization Problem:

(1) Write $q(x)=x^TSx$ for S symmetric

(2) Orthogonally diagonalize S: orthonormal eigenbasis

$$S=QDQ^T \quad Q=\begin{pmatrix}| & & |\\ \omega_1 & \cdots & \omega_n\\ | & & |\end{pmatrix} \text{ orthogonal},\quad D=\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}$$

Order so that $\lambda_1\geq\cdots\geq\lambda_n$.

$\to$ Set $y=Q^Tx$. Then
$$q(x)=x^TQDQ^Tx=(Q^Tx)D(Q^Tx)=y^TDy$$
$$=\lambda_1y_1^2+\cdots+\lambda_ny_n^2$$

q is diagonal in the y-coordinates.

(3) The maximum value of q is the largest eigenvalue λ_1:

$q(\omega_1)=\lambda_1$ b/c $Q^T\omega_1=\begin{pmatrix}1\\ \vdots\\ 0\end{pmatrix}$ ie. $y_1=1,\ y_2=\cdots=y_n=0$

The minimum value of q is the smallest eigenvalue λ_n:

$q(\omega_n)=\lambda_n$ b/c $Q^T\omega_n=\begin{pmatrix}0\\ \vdots\\ 1\end{pmatrix}$ ie $y_1=\cdots=y_{n-1}=0,\ y_n=1$

Eg: (from last time) $q(x_1,x_2)=\frac{5}{2}(x_1^2+x_2^2)-x_1x_2$

$$q(x)=x^TSx \qquad S=\frac{1}{2}\begin{pmatrix}5 & -1\\ -1 & 5\end{pmatrix}=Q^TDQ$$

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

$\Rightarrow$ max is $q(1/\sqrt{2}, -1/\sqrt{2})=3$

min is $q(1/\sqrt{2}, 1/\sqrt{2})=2$

Ellipsoids

Suppose all $\lambda_i>0$. Then

$q(x)=\lambda_1y_1^2+\cdots+\lambda_ny_n^2=1$ is an ellipsoid.

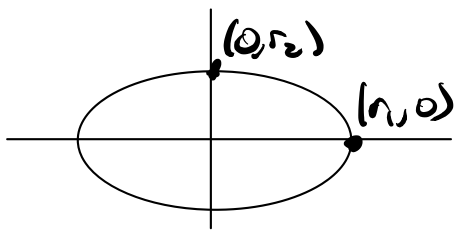
If you take the $(n-1)$-sphere

$$x_1^2+\cdots+x_n^2=1$$

and change coords $y_i=\sqrt{\lambda_i}\,x_i$.

When $n=2$ you get an ellipse:

$$\left(\frac{x_1}{r_1}\right)^2+\left(\frac{x_2}{r_2}\right)^2=1 \rightsquigarrow$$

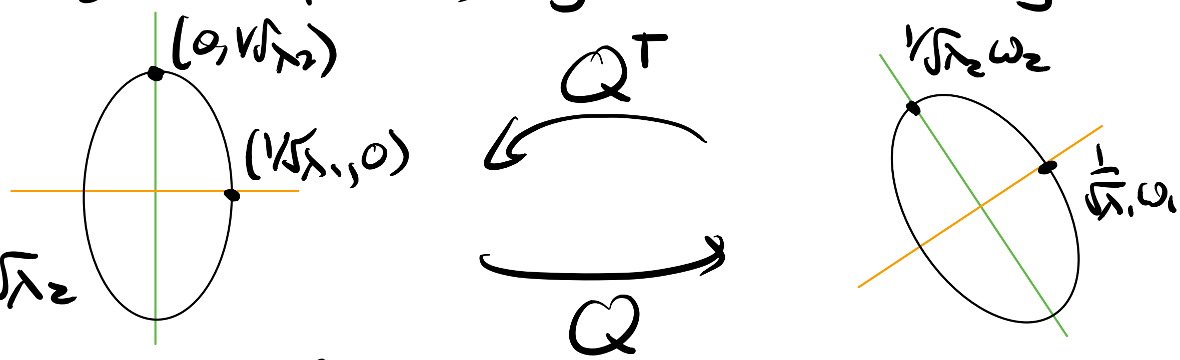


When you diagonalize $q(x_1,x_2)$, $y=Q^Tx \Leftrightarrow x=Qy$

$\lambda_1\geq\lambda_2$

$\frac{1}{\sqrt{\lambda_1}}\leq\frac{1}{\sqrt{\lambda_2}}$

$r_1=1/\sqrt{\lambda_1}$ $r_2=1/\sqrt{\lambda_2}$



$\lambda_1y_1^2+\lambda_2y_2^2=1$ $\qquad$ $q(x_1,x_2)=1$

Relation to Quadratic Optimization:

Minimize $\|x\|$ subject to $q(x)=1 \rightsquigarrow x=1/\sqrt{\lambda_1}\,\omega_1$

$\Rightarrow q(\omega_1)=q(x/\|x\|)=1/\|x\|^2\,\cancelto{1}{q(x)}=\lambda_1$ maximized.

Maximize $\|x\|$ subject to $q(x)=1 \rightsquigarrow x = \frac{1}{\sqrt{\lambda_2}}\,\omega_2$

$$\Rightarrow q(\omega_2) = q\left(\frac{x}{\|x\|}\right) = \frac{1}{\|x\|^2}\cancel{q(x)}^{1} = \lambda_2 \quad \text{minimized.}$$

Positive-Definite matrices

The condition "all $\lambda_i > 0$" is special — and very common.

Def: A symmetric matrix is

- positive-definite if all eigenvalues are >0
- positive-semidefinite if all eigenvalues are ≥ 0
- Likewise for negative-definite & negative-semidefinite.
- Indefinite otherwise.

+ve-def	+ve-semidef	−ve-def	indefinite
$Q\begin{pmatrix}3&0\\0&2\end{pmatrix}Q^T$	$Q\begin{pmatrix}3&0\\0&0\end{pmatrix}Q^T$	$Q\begin{pmatrix}-2&0\\0&-3\end{pmatrix}Q^T$	$Q\begin{pmatrix}3&0\\0&-2\end{pmatrix}Q^T$

Positive-definiteness is an important condition! Want to check without computing eigenvalues!

Conditions for S to be Positive-Definite:

The following are equivalent for a symmetric matrix S:

(1) All eigenvalues are >0 — +ve-def

(2) $x^T S x > 0$ for all $x \neq 0$ — +ve-energy

(3) All n upper-left determinants are >0.

$$S = \left(\begin{array}{c|c} S_1 & \text{col} \\ \hline \text{row} & \end{array}\right) \qquad \begin{array}{l} \det(S) > 0 \\ \det(S_1) > 0 \end{array} \quad \text{etc.}$$

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

(5) $S = LU$ (no row swaps) and diagonal entries of U are >0. $\leftarrow$ fastest! (elimination problem)

Eg: $S = \begin{pmatrix} 7 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 5 \end{pmatrix}$ $\quad p(\lambda) = -(\lambda-3)(\lambda-6)(\lambda-9)$

(1) Eigenvalues are $3, 6, 9 > 0$.

(2) $x^TSx = 9y_1^2 + 6y_2^2 + 3y_3^2 > 0$ $\quad S = QDQ^T \quad y = Q^Tx$

1×1 matrix whenever $x \neq 0$ $(\Rightarrow Q^Tx = y \neq 0)$

(3) $\det(7) = 7 > 0 \quad \det\begin{pmatrix} 7 & 2 \\ 2 & 6 \end{pmatrix} = 38 > 0 \quad \det(S) = 3\cdot 6\cdot 9 > 0$

(4) Cholesky: next

(5) $\begin{pmatrix} 7 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 5 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 7 & 2 & 0 \\ 0 & 38/7 & 2 \\ 0 & 2 & 5 \end{pmatrix} \longrightarrow \begin{pmatrix} 7 & 2 & 0 \\ 0 & 38/7 & 2 \\ 0 & 0 & 4\frac{10}{38} \end{pmatrix} = U$

↑ diag >0

Whence these conditions?

(1) is the definition

(2) $x^TSx = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$. All $\lambda_i > 0 \Rightarrow$

this is >0 when $x \neq 0$ $(\Rightarrow y = Q^Tx \neq 0)$

(3) Determinants: magic

(4) If $S = A^TA$ for A with full col rank:

$x^TSx = x^TA^TAx = (Ax)^T(Ax) = (Ax)\cdot(Ax) = \|Ax\|^2$

If A has FCR then $x \neq 0 \Rightarrow Ax \neq 0$.

(5) From $S = LDL^T$: next.

LDL^T and Cholesky.

Basically LU for +ve-def symmetric matrices.

Thm: A positive-definite symmetric matrix S can be uniquely decomposed as

$$S = LDL^T \quad \text{and} \quad S = L_1 L_1^T \leftarrow \text{Cholesky}$$

where:

- D: diagonal w/ positive diagonal entries
- L: lower-unitriangular
- L_1: lower-triangular with positive diagonal entries.

Pf: [supplement]

NB: L_1^T has full col rank (it's invertible) and $S = A^TA$ for $A = L_1^T$. So this justifies condition 3.

NB: If $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$ set $\sqrt{D} = \begin{pmatrix} \sqrt{d_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{d_n} \end{pmatrix}$

Define $L_1 = L\sqrt{D}$ (scales cols of L by $\sqrt{d_1}, \dots, \sqrt{d_n}$)

Then $L_1 L_1^T = L\sqrt{D}\sqrt{D}^T L^T = LDL^T = S$.

So LDL^T & $L_1 L_1^T$ are interchangeable.

We'll focus on LDL^T.

NB: If $S = LDL^T$ set $U = DL^T$ (scales rows of L^T by entries of D). Then $S = LU$ is the LU decomposition!

Procedure to compute $S = LDL^T$: S symmetric, +ve-def

(1) Find $S = LU$

Magic: never need row swaps!

(2) $L = L$ $\quad D$ = diagonal entries of U: $S = LDL^T$

Magic: Diagonal entries of U are > 0.

Magic: $DL^T = U$ i.e. $D^{-1}U = L^T$.

Why LDL^T / $L_1 L_1^T$?

- Solve $Sx = b$ by forward/back substitution (same reason you want LU)
- Can compute LDL^T about twice as fast as LU, & with half the memory.

  LU: $\frac{2}{3}n^3$ flops $\qquad LDL^T$: $\frac{1}{3}n^3$ flops

  [supplement]