Symmetric Positive Definite Matrices

Zed

November 19, 2016

1 Preliminaries

1.1 Inner Products

- Inner product on $\mathbb{R}^n$: $\langle \boldsymbol{v}, \boldsymbol{u} \rangle = \boldsymbol{v}^\top \boldsymbol{u}$. Has 3 properties:
  1. Positivity: $\langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0$; $\langle \boldsymbol{v}, \boldsymbol{v} \rangle = 0 \iff \boldsymbol{v} = \boldsymbol{0}$.
  2. Bilinearity: $\langle a\boldsymbol{x} + b\boldsymbol{y}, \boldsymbol{z} \rangle = a\langle \boldsymbol{x}, \boldsymbol{z} \rangle + b\langle \boldsymbol{y}, \boldsymbol{z} \rangle$; $\langle \boldsymbol{z}, a\boldsymbol{x} + b\boldsymbol{y} \rangle = a\langle \boldsymbol{z}, \boldsymbol{x} \rangle + b\langle \boldsymbol{z}, \boldsymbol{y} \rangle$.
  3. Symmetry: $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$.
- Norm on $\mathbb{R}^n$: $\|\boldsymbol{v}\|^2 = \langle \boldsymbol{v}, \boldsymbol{v} \rangle$
- Inner product on $\mathbb{C}^n$: $\langle \boldsymbol{v}, \boldsymbol{u} \rangle_{\mathbb{C}} = \boldsymbol{v}^H \boldsymbol{u}$. Where $\boldsymbol{v}^H = (\bar{v}_1, ..., \bar{v}_n)$ is conjugate transpose of col vector $\boldsymbol{v}$, $\bar{v}$ is complex conjugate of entry v, i.e.

$$\langle \boldsymbol{v}, \boldsymbol{u} \rangle_{\mathbb{C}} = \sum_{j=1}^{n} u_j \bar{v}_j$$

  Also 3 properties:
  1. Positivity: $\langle \boldsymbol{v}, \boldsymbol{v} \rangle_{\mathbb{C}} \geq 0$; $\langle \boldsymbol{v}, \boldsymbol{v} \rangle_{\mathbb{C}} = 0 \iff \boldsymbol{v} = \boldsymbol{0}$.
  2. Sesquilinearity:
  $$\langle a\boldsymbol{x} + b\boldsymbol{y}, \boldsymbol{z} \rangle_{\mathbb{C}} = a\langle \boldsymbol{x}, \boldsymbol{z} \rangle_{\mathbb{C}} + b\langle \boldsymbol{y}, \boldsymbol{z} \rangle_{\mathbb{C}}$$
  $$\langle \boldsymbol{z}, a\boldsymbol{x} + b\boldsymbol{y} \rangle_{\mathbb{C}} = \bar{a}\langle \boldsymbol{z}, \boldsymbol{x} \rangle_{\mathbb{C}} + \bar{b}\langle \boldsymbol{z}, \boldsymbol{y} \rangle_{\mathbb{C}}$$
     Proof. Use conjugate symmetry. $\langle \boldsymbol{z}, a\boldsymbol{x} + b\boldsymbol{y} \rangle_{\mathbb{C}} = \overline{\langle a\boldsymbol{x} + b\boldsymbol{y}, \boldsymbol{z} \rangle_{\mathbb{C}}} = \bar{a}\overline{\langle \boldsymbol{x}, \boldsymbol{z} \rangle_{\mathbb{C}}} + \bar{b}\overline{\langle \boldsymbol{y}, \boldsymbol{z} \rangle_{\mathbb{C}}} = \bar{a}\langle \boldsymbol{z}, \boldsymbol{x} \rangle_{\mathbb{C}} + \bar{b}\langle \boldsymbol{z}, \boldsymbol{y} \rangle$
  3. Conjugate Symmetry: $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle}$.
- Norm on $\mathbb{C}^n$: $\|v\|_{\mathbb{C}}^2 = \langle \boldsymbol{v}, \boldsymbol{v} \rangle_{\mathbb{C}} = \boldsymbol{v}^H \boldsymbol{v}$.
- Let $\boldsymbol{A}$ be a matrix with complex entries, its conjugate transpose (hermitian): $\boldsymbol{A}^H = \overline{\boldsymbol{A}}^\top$. We have $(\boldsymbol{A}\boldsymbol{B})^H = \boldsymbol{B}^H \boldsymbol{A}^H$.
  And by definition of inner product on complex field, $\langle \boldsymbol{A}\boldsymbol{u}, \boldsymbol{v} \rangle_{\mathbb{C}} = \boldsymbol{v}^H \boldsymbol{A}\boldsymbol{u} = \langle \boldsymbol{u}, (\boldsymbol{v}^H \boldsymbol{A})^H \rangle_{\mathbb{C}} = \langle \boldsymbol{u}, \boldsymbol{A}^H \boldsymbol{v} \rangle_{\mathbb{C}}$. Similarly $\langle \boldsymbol{u}, \boldsymbol{B}\boldsymbol{v} \rangle = \langle \boldsymbol{B}^H \boldsymbol{u}, \boldsymbol{v} \rangle$.

2 Symmetric and Orthogonal Matrix

2.1 Basics

- Symmetric matrix: $\boldsymbol{A} = \boldsymbol{A}^\top$. Let $\boldsymbol{X}$ be an arbitrary matrix, $\boldsymbol{X}^\top \boldsymbol{X}$ and $(\boldsymbol{X} + \boldsymbol{X}^\top)$ are symmetric.

· $\langle \boldsymbol{Au}, \boldsymbol{v} \rangle = \boldsymbol{v}^\top \boldsymbol{Au} = \langle \boldsymbol{u}, \boldsymbol{A}^\top \boldsymbol{v} \rangle$. And $\langle \boldsymbol{u}, \boldsymbol{Bv} \rangle = \boldsymbol{u}^\top \boldsymbol{Bv} = \langle \boldsymbol{B}^\top \boldsymbol{u}, \boldsymbol{v} \rangle$.

· $\boldsymbol{u} \perp \boldsymbol{v} \iff \langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$.

· A matrix $\boldsymbol{Q}$ is orthogonal iff any two different columns of it are orthonormal. (orthogonal and unit norm); or any two different rows of it are orthonormal.

Thm. A square matrix $\boldsymbol{Q}$ is orthogonal iff $\boldsymbol{Q}^{-1} = \boldsymbol{Q}^\top$.
Proof. We have $\boldsymbol{Q}^\top \boldsymbol{Q} = \boldsymbol{I}$.

$$\boldsymbol{Q}^\top \boldsymbol{Q} = \begin{pmatrix} \boldsymbol{q}_1^\top \\ \boldsymbol{q}_2^\top \\ \vdots \\ \boldsymbol{q}_n^\top \end{pmatrix} \begin{pmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 & \cdots & \boldsymbol{q}_n \end{pmatrix} = \begin{pmatrix} \|\boldsymbol{q}_1\|^2 & \langle \boldsymbol{q}_1, \boldsymbol{q}_2 \rangle & \cdots & \langle \boldsymbol{q}_1, \boldsymbol{q}_n \rangle \\ \langle \boldsymbol{q}_2, \boldsymbol{q}_1 \rangle & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \langle \boldsymbol{q}_n, \boldsymbol{q}_1 \rangle & \cdots & \cdots & \|\boldsymbol{q}_n\|^2 \end{pmatrix}$$

Hence $\boldsymbol{Q}^\top \boldsymbol{Q} = \boldsymbol{I} \iff \|\boldsymbol{q}_i\| = 1$ for all i and $\langle \boldsymbol{q}_j, \boldsymbol{q}_k \rangle = 0$ for $k \neq j$. □

· Properties of $\boldsymbol{Q}$:

1. If $\boldsymbol{Q}_1, \boldsymbol{Q}_2$ orthogonal matrices, same size $\Rightarrow \boldsymbol{Q}_1 \boldsymbol{Q}_2$ orthogonal.
2. $\boldsymbol{v}$ is $n \times 1$ vector, then $\|\boldsymbol{v}\| = \|\boldsymbol{Qv}\|$.
3. If λ is eigval of $\boldsymbol{Q} \Rightarrow |\lambda| = 1$.

Proof. First prop: $\boldsymbol{Q}_1 \boldsymbol{Q}_2 (\boldsymbol{Q}_1 \boldsymbol{Q}_2)^\top = \boldsymbol{Q}_1 \boldsymbol{Q}_2 \boldsymbol{Q}_2^\top \boldsymbol{Q}_1^\top = \boldsymbol{I}$.
Second: $\langle \boldsymbol{Qv}, \boldsymbol{Qv} \rangle = (\boldsymbol{Qv})^\top \boldsymbol{Qv} = \boldsymbol{v}^\top \boldsymbol{Q}^\top \boldsymbol{Qv} = \boldsymbol{v}^\top \boldsymbol{v} = \|\boldsymbol{v}\|^2$.
Third: by (2), $\|\boldsymbol{v}\| = \|\boldsymbol{Qv}\| = \|\lambda \boldsymbol{v}\| = |\lambda| \, \|\boldsymbol{v}\| \Rightarrow |\lambda| = 1$.

2.2 Eigvals and Eigvecs

Thm. Any eigval of symmetric matrix is real number.
Proof. We have $\boldsymbol{Av} = \lambda \boldsymbol{v}$. $\boldsymbol{A}$ symmetric and real, hence $\boldsymbol{A} = \boldsymbol{A}^H$. Consider

$$\langle \boldsymbol{Av}, \boldsymbol{v} \rangle_{\mathbb{C}} = \boldsymbol{v}^H \boldsymbol{Av} = \langle \boldsymbol{v}, \boldsymbol{A}^H \boldsymbol{v} \rangle_{\mathbb{C}} = \langle \boldsymbol{v}, \boldsymbol{Av} \rangle_{\mathbb{C}}$$

And $\langle \boldsymbol{Av}, \boldsymbol{v} \rangle_{\mathbb{C}} = \langle \lambda \boldsymbol{v}, \boldsymbol{v} \rangle_{\mathbb{C}} = \lambda \|\boldsymbol{v}\|_{\mathbb{C}}^2$; $\langle \boldsymbol{v}, \boldsymbol{Av} \rangle_{\mathbb{C}} = \langle \boldsymbol{v}, \lambda \boldsymbol{v} \rangle_{\mathbb{C}} = \bar{\lambda} \|\boldsymbol{v}\|_{\mathbb{C}}^2$
$\Rightarrow \lambda = \bar{\lambda}$, which implies that λ is real. □

· Eigvecs corresponding to different eigvals of symmetric matrix are orthogonal.
Proof. $\langle \boldsymbol{Av}_1, \boldsymbol{v}_2 \rangle = \langle \lambda_1 \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = \lambda_1 \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle$.
$\langle \boldsymbol{Av}_1, \boldsymbol{v}_2 \rangle = \langle \boldsymbol{v}_1, \boldsymbol{A}^\top \boldsymbol{v}_2 \rangle = \langle \boldsymbol{v}_1, \boldsymbol{Av}_2 \rangle = \langle \boldsymbol{v}_1, \lambda_2 \boldsymbol{v}_2 \rangle = \lambda_2 \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle$.
Hence $\lambda_2 \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = \lambda_1 \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle$; $\lambda_1 \neq \lambda_2 \Rightarrow \boldsymbol{v}_1 \perp \boldsymbol{v}_2$. □

2.3 Diagonal Form

Thm. $\boldsymbol{A}$ is symmetric matrix, then $\boldsymbol{A}$ is diagonalizable, $\boldsymbol{A} = \boldsymbol{Q} \boldsymbol{\Lambda} \boldsymbol{Q}^{-1}$. Its eigvals are entries of $\boldsymbol{\Lambda}$, cols of $\boldsymbol{Q}$ are eigvecs, and $\boldsymbol{Q}$ is orthogonal.
So $\boldsymbol{Q}^\top = \boldsymbol{Q}^{-1}$, we can also write $\boldsymbol{A} = \boldsymbol{Q} \boldsymbol{\Lambda} \boldsymbol{Q}^\top$.
Proof. $\boldsymbol{A}^{1 \times 1}$ case is trivial. We wanna prove by induction.
Assume $\boldsymbol{A}_{n-1} = \boldsymbol{Q}_{n-1} \boldsymbol{\Lambda}_{n-1} \boldsymbol{Q}_{n-1}^{-1}$ with shape $n-1 \times n-1$.
Now consider $\boldsymbol{A}^{n \times n}$. $(\lambda_1, \boldsymbol{v}_1)$ being an eigentuple of $\boldsymbol{A}$, pick $\|\boldsymbol{v}_1\| = 1$. Construct $\boldsymbol{Q}_1 =$

$(\boldsymbol{v}_1, \boldsymbol{q}_2, ..., \boldsymbol{q}_n)$ such that it is orthogonal. We have

$$\begin{aligned}\boldsymbol{Q}_1^\top \boldsymbol{A}\boldsymbol{Q}_1 &= \begin{pmatrix}\boldsymbol{v}_1^\top \\ \boldsymbol{q}_2^\top \\ \vdots \\ \boldsymbol{q}_n^\top\end{pmatrix} \begin{pmatrix}\boldsymbol{A}\boldsymbol{v}_1 & \boldsymbol{A}\boldsymbol{q}_2 & \cdots & \boldsymbol{A}\boldsymbol{q}_n\end{pmatrix} \\ &= \begin{pmatrix}\lambda_1 \boldsymbol{v}_1^\top \boldsymbol{v}_1 & \boldsymbol{v}_1^\top \boldsymbol{A}\boldsymbol{q}_2 & \cdots & \boldsymbol{v}_1^\top \boldsymbol{A}\boldsymbol{q}_n \\ \lambda_1 \boldsymbol{q}_2^\top \boldsymbol{v}_1 & \boldsymbol{q}_2^\top \boldsymbol{A}\boldsymbol{q}_2 & \cdots & \boldsymbol{q}_2^\top \boldsymbol{A}\boldsymbol{q}_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 \boldsymbol{q}_n^\top \boldsymbol{v}_1 & \boldsymbol{q}_n^\top \boldsymbol{A}\boldsymbol{q}_2 & \cdots & \boldsymbol{q}_n^\top \boldsymbol{A}\boldsymbol{q}_n\end{pmatrix} \\ &= \begin{pmatrix}\lambda_1 & \boldsymbol{v}_1^\top \boldsymbol{A}\boldsymbol{q}_2 & \cdots & \boldsymbol{v}_1^\top \boldsymbol{A}\boldsymbol{q}_n \\ 0 & \boldsymbol{q}_2^\top \boldsymbol{A}\boldsymbol{q}_2 & \cdots & \boldsymbol{q}_2^\top \boldsymbol{A}\boldsymbol{q}_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \boldsymbol{q}_n^\top \boldsymbol{A}\boldsymbol{q}_2 & \cdots & \boldsymbol{q}_n^\top \boldsymbol{A}\boldsymbol{q}_n\end{pmatrix}\end{aligned}$$

$\boldsymbol{Q}_1^\top \boldsymbol{A}\boldsymbol{Q}_1$ is symmetric, so other entries on first row are also zeros. The southeast $n-1 \times n-1$ block is $\boldsymbol{A}_{n-1}$, by assumption it can be written as $\boldsymbol{A}_{n-1} = \boldsymbol{Q}_{n-1}\boldsymbol{\Lambda}_{n-1}\boldsymbol{Q}_{n-1}^{-1} = \boldsymbol{Q}_{n-1}\boldsymbol{\Lambda}_{n-1}\boldsymbol{Q}_{n-1}^t$, $\boldsymbol{Q}_{n-1}$ orthogonal, Λ_{n-1} diagonal. So

$$\begin{aligned}\boldsymbol{Q}_1^\top \boldsymbol{A}\boldsymbol{Q}_1 &= \begin{pmatrix}\lambda_1 & \boldsymbol{0} \\ \boldsymbol{0}^\top & \boldsymbol{Q}_{n-1}\boldsymbol{\Lambda}_{n-1}\boldsymbol{Q}_{n-1}^t\end{pmatrix} \\ &= \begin{pmatrix}1 & \boldsymbol{0} \\ \boldsymbol{0}^\top & \boldsymbol{Q}_{n-1}\end{pmatrix}\begin{pmatrix}\lambda_1 & \boldsymbol{0} \\ \boldsymbol{0}^\top & \boldsymbol{\Lambda}_{n-1}\end{pmatrix}\begin{pmatrix}1 & \boldsymbol{0} \\ \boldsymbol{0}^\top & \boldsymbol{Q}_{n-1}^\top\end{pmatrix} \\ &= \boldsymbol{Q}_n\boldsymbol{\Lambda}_n\boldsymbol{Q}_n^\top\end{aligned}$$

$\Rightarrow \boldsymbol{A} = \boldsymbol{Q_1}\boldsymbol{Q}_n\boldsymbol{\Lambda}_n\boldsymbol{Q}_n^\top\boldsymbol{Q}_1^\top = (\boldsymbol{Q_1}\boldsymbol{Q}_n)\boldsymbol{\Lambda}_n(\boldsymbol{Q_1}\boldsymbol{Q}_n)^\top$. Let $\boldsymbol{Q} := \boldsymbol{Q}_1\boldsymbol{Q}_n$, it's also orthogonal by prop 1. Hence $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}_n\boldsymbol{Q}^\top = \boldsymbol{Q}\boldsymbol{\Lambda}_n\boldsymbol{Q}^{-1}$. $\square$

3 Symmetric Positive Definite Matrix

· $\boldsymbol{A}$ is symmetric positive definite (spd) iff

$$\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}\rangle = \boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x} > 0, \ \ \forall \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x} \neq \boldsymbol{0}.$$

This def (spd) is equivalent to

$$\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x} \geq 0, \ \ \forall \boldsymbol{x} \in \mathbb{R}^n \quad \text{and} \quad \boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x} = 0 \iff \boldsymbol{x} = 0$$

$\boldsymbol{A}$ is symmetric positive semidefinite (spsd) iff $\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}\rangle \geq 0$, $\forall \boldsymbol{x} \in \mathbb{R}^n$.
$\boldsymbol{A}$ is symmetric *negative* definite iff $\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}\rangle < 0$, $\forall x \in \mathbb{R}^n$. Iff $-\boldsymbol{A}$ is spd.
$\boldsymbol{A}$ is symmetric *negative* semidefinite iff $-\boldsymbol{A}$ is spsd.

· $\boldsymbol{M}^\top\boldsymbol{M}$ is spsd, it is spd iff $\boldsymbol{M}$ has full rank, i.e. nonsingular.
Proof. $\langle \boldsymbol{M}^\top\boldsymbol{M}\boldsymbol{x}, \boldsymbol{x}\rangle = \langle \boldsymbol{M}\boldsymbol{x}, \boldsymbol{M}\boldsymbol{x}\rangle = \|\boldsymbol{M}\boldsymbol{x}\|^2 \geq 0$.
Moreover, $\|\boldsymbol{M}\boldsymbol{x}\|^2 = 0 \iff \boldsymbol{M}\boldsymbol{x} = 0$. If $\boldsymbol{x} \neq 0$; $\boldsymbol{M}\boldsymbol{x} \neq 0 \iff \boldsymbol{M}$ has full rank, because $\boldsymbol{M}\boldsymbol{x}$ is just linear comb of cols of $\boldsymbol{M}$ with weights $\boldsymbol{x}$. $\square$

Thm. Symmetic matrix $\boldsymbol{A}$ is spd $\iff$ all eigvals of it are **strictly** greater than 0.
It's spsd $\iff$ all eigvals are greater than or equal to 0.
Proof. We show the spsd case.

($\Leftarrow$): By diagonal form of symmetric matrix: $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^\top$.
So $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^\top \boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{\Lambda}\boldsymbol{y}$, where $\boldsymbol{y} = \boldsymbol{Q}^\top \boldsymbol{x}$, i.e.

$$\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{\Lambda}\boldsymbol{y} = \sum_{j=1}^{n} \lambda_j y_j^2 \geq 0, \quad \forall \boldsymbol{y} \in \mathbb{R}^n, \boldsymbol{y} \neq \boldsymbol{0}.$$

since all $\lambda_j \geq 0$. So $\boldsymbol{A}$ is spsd.
($\Rightarrow$): Suppose there is a $\lambda < 0$, with eigvec $\boldsymbol{v}$. Then

$$\langle \boldsymbol{A}\boldsymbol{v}, \boldsymbol{v} \rangle = \lambda \|\boldsymbol{v}\|^2 < 0$$

contradicts the fact that $\boldsymbol{A}$ is spsd. Similar proof for strictly positive eigvals and spd case.
$\square$

- Spd matrix is nonsingular (due to nonzero eigvals). The inverse of spd matrix is also spd. (eigvals are $\frac{1}{\lambda} > 0$.)
- symmetric + strictly diagonally dominant + positive entries on main diag $\Rightarrow$ spd.
  Symmetric + Weakly diagonally dominant + positive entries on main diag $\Rightarrow$ spsd.
  Proof. By (Gershgorin): $|\lambda - A_{jj}| \leq R_j$. $A_{jj} - \lambda \leq |A_{jj} - \lambda| \leq R_j$
  $\Rightarrow \lambda \geq A_{jj} - R_j$. If $\boldsymbol{A}$ is strictly diagdom with positive diag entries: $A_{jj} = |A_{jj}| > R_j$.
  Hence $\lambda \geq |A_{jj}| - R_j > 0$. All its eigvals are positive $\Rightarrow$ spd. $\square$

Thm. (**Sylvester's Criterion**) A symmetric matrix is spd $\iff$ all its leading principal minors are positive.
It is spsd $\iff$ all its principal minors greater than or equal to 0.