

Linear Control and Estimation

Positive Definiteness and Matrix Norm

Sivakumar Balasubramanian

Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

References

- ▶ G Strang, Linear Algebra: Chapters 6 and 7.

Positive definite matrices

- ▶ We know that $\mathbf{x}^T \mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{x} \neq \mathbf{0}$.
- ▶ What can we say about $\mathbf{x}^T \mathbf{A} \mathbf{x}$, where $\mathbf{A} \in \mathbb{R}^{n \times n}$?
 - ▶ Is it positive for all $\mathbf{x} \neq \mathbf{0}$?
 - ▶ Can it be zero for some $\mathbf{x} \neq \mathbf{0}$?
 - ▶ Can it be negative some $\mathbf{x} \neq \mathbf{0}$?
- ▶ Any matrix \mathbf{A} for which $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$, for all $\mathbf{x} \neq \mathbf{0}$ is called a *positive definite matrix*.
- ▶ Positive definite matrices are very useful and are commonly encountered very in practice: optimization, mechanics (mass matrix, stiffness matrix), stability analysis, co-variance matrices etc.

Positive definite matrices

- ▶ We know that $\mathbf{x}^T \mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{x} \neq \mathbf{0}$.
- ▶ What can we say about $\mathbf{x}^T \mathbf{A} \mathbf{x}$, where $\mathbf{A} \in \mathbb{R}^{n \times n}$?
 - ▶ Is it positive for all $\mathbf{x} \neq \mathbf{0}$? E.g. \mathbf{I} , $\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$
 - ▶ Can it be zero for some $\mathbf{x} \neq \mathbf{0}$?
 - ▶ Can it be negative some $\mathbf{x} \neq \mathbf{0}$?
- ▶ Any matrix \mathbf{A} for which $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$, for all $\mathbf{x} \neq \mathbf{0}$ is called a *positive definite matrix*.
- ▶ Positive definite matrices are very useful and are commonly encountered very in practice: optimization, mechanics (mass matrix, stiffness matrix), stability analysis, co-variance matrices etc.

Positive definite matrices

- ▶ We know that $\mathbf{x}^T \mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{x} \neq \mathbf{0}$.
- ▶ What can we say about $\mathbf{x}^T \mathbf{A} \mathbf{x}$, where $\mathbf{A} \in \mathbb{R}^{n \times n}$?
 - ▶ Is it positive for all $\mathbf{x} \neq \mathbf{0}$? E.g. \mathbf{I} , $\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$
 - ▶ Can it be zero for some $\mathbf{x} \neq \mathbf{0}$? E.g. $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$
 - ▶ Can it be negative some $\mathbf{x} \neq \mathbf{0}$?
- ▶ Any matrix \mathbf{A} for which $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$, for all $\mathbf{x} \neq \mathbf{0}$ is called a *positive definite matrix*.
- ▶ Positive definite matrices are very useful and are commonly encountered very in practice: optimization, mechanics (mass matrix, stiffness matrix), stability analysis, co-variance matrices etc.

Positive definite matrices

- ▶ We know that $\mathbf{x}^T \mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{x} \neq \mathbf{0}$.
- ▶ What can we say about $\mathbf{x}^T \mathbf{A} \mathbf{x}$, where $\mathbf{A} \in \mathbb{R}^{n \times n}$?
 - ▶ Is it positive for all $\mathbf{x} \neq \mathbf{0}$? E.g. \mathbf{I} , $\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$
 - ▶ Can it be zero for some $\mathbf{x} \neq \mathbf{0}$? E.g. $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$
 - ▶ Can it be negative some $\mathbf{x} \neq \mathbf{0}$? E.g. $\begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}$, $\begin{bmatrix} 1 & -4 \\ -4 & 1 \end{bmatrix}$
- ▶ Any matrix \mathbf{A} for which $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$, for all $\mathbf{x} \neq \mathbf{0}$ is called a *positive definite matrix*.
- ▶ Positive definite matrices are very useful and are commonly encountered very in practice: optimization, mechanics (mass matrix, stiffness matrix), stability analysis, co-variance matrices etc.

Positive definite matrices

- ▶ We know that $\mathbf{x}^T \mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{x} \neq \mathbf{0}$.
- ▶ What can we say about $\mathbf{x}^T \mathbf{A} \mathbf{x}$, where $\mathbf{A} \in \mathbb{R}^{n \times n}$?
 - ▶ Is it positive for all $\mathbf{x} \neq \mathbf{0}$? E.g. \mathbf{I} , $\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$
 - ▶ Can it be zero for some $\mathbf{x} \neq \mathbf{0}$? E.g. $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$
 - ▶ Can it be negative some $\mathbf{x} \neq \mathbf{0}$? E.g. $\begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}$, $\begin{bmatrix} 1 & -4 \\ -4 & 1 \end{bmatrix}$
- ▶ Any matrix \mathbf{A} for which $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$, for all $\mathbf{x} \neq \mathbf{0}$ is called a *positive definite matrix*.
- ▶ Positive definite matrices are very useful and are commonly encountered very in practice: optimization, mechanics (mass matrix, stiffness matrix), stability analysis, co-variance matrices etc.

Are the following matrices positive definite: $\begin{bmatrix} 2 & -6 \\ 2 & 1 \end{bmatrix}$, $\begin{bmatrix} 5 & -1 \\ 14 & 11 \end{bmatrix}$

Positive Definite Matrix

- ▶ The idea of positive definiteness is intimately related to the problem of minimization of a function.
- ▶ Consider the following function of a single variable $f(x)$. This function reaches a minimum at $x = 0$, when $\frac{df(x)}{dx} \big|_{x=0} = 0$ and $\frac{d^2f(x)}{dx^2} \big|_{x=0} > 0$. E.g.,

$$f(x) = 3x^2 \rightarrow \frac{df(x)}{dx} \big|_{x=0} = 0, \frac{d^2f(x)}{dx^2} \big|_{x=0} = 3 > 0$$

- ▶ What about $f(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$? We can extend the previous idea using partial derivatives.

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = 0, \frac{\partial f(x_1, x_2)}{\partial x_2} = 0, \frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} > 0, \frac{\partial^2 f(x_1, x_2)}{\partial x_2^2} > 0$$

Is this enough? $\frac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_2}$ must also be taken into account.

Positive Definite Matrix

- ▶ The idea of positive definiteness is intimately related to the problem of minimization of a function.
- ▶ Consider the following function of a single variable $f(x)$. This function reaches a minimum at $x = 0$, when $\frac{df(x)}{dx}\big|_{x=0} = 0$ and $\frac{d^2f(x)}{dx^2}\big|_{x=0} > 0$. E.g.,

$$f(x) = 3x^2 \rightarrow \frac{df(x)}{dx}\big|_{x=0} = 0, \quad \frac{d^2f(x)}{dx^2}\big|_{x=0} = 3 > 0$$

- ▶ What about $f(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$? We can extend the previous idea using partial derivatives.

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = 0, \quad \frac{\partial f(x_1, x_2)}{\partial x_2} = 0, \quad \frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} > 0, \quad \frac{\partial^2 f(x_1, x_2)}{\partial x_2^2} > 0$$

Is this enough? $\frac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_2}$ must also be taken into account.

Are these functions positive for all x_1, x_2 ? 1) $x_1^2 + x_1x_2 + x_2^2$,

Positive Definite Matrix

- ▶ The idea of positive definiteness is intimately related to the problem of minimization of a function.
- ▶ Consider the following function of a single variable $f(x)$. This function reaches a minimum at $x = 0$, when $\frac{df(x)}{dx}\big|_{x=0} = 0$ and $\frac{d^2f(x)}{dx^2}\big|_{x=0} > 0$. E.g.,

$$f(x) = 3x^2 \rightarrow \frac{df(x)}{dx}\big|_{x=0} = 0, \quad \frac{d^2f(x)}{dx^2}\big|_{x=0} = 3 > 0$$

- ▶ What about $f(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$? We can extend the previous idea using partial derivatives.

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = 0, \quad \frac{\partial f(x_1, x_2)}{\partial x_2} = 0, \quad \frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} > 0, \quad \frac{\partial^2 f(x_1, x_2)}{\partial x_2^2} > 0$$

Is this enough? $\frac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_2}$ must also be taken into account.

Are these functions positive for all x_1, x_2 ? 1) $x_1^2 + x_1x_2 + x_2^2$, 2) $x_1^2 + 2x_1x_2 + x_2^2$,

Positive Definite Matrix

- ▶ The idea of positive definiteness is intimately related to the problem of minimization of a function.
- ▶ Consider the following function of a single variable $f(x)$. This function reaches a minimum at $x = 0$, when $\frac{df(x)}{dx}\big|_{x=0} = 0$ and $\frac{d^2f(x)}{dx^2}\big|_{x=0} > 0$. E.g.,

$$f(x) = 3x^2 \rightarrow \frac{df(x)}{dx}\big|_{x=0} = 0, \quad \frac{d^2f(x)}{dx^2}\big|_{x=0} = 3 > 0$$

- ▶ What about $f(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$? We can extend the previous idea using partial derivatives.

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = 0, \quad \frac{\partial f(x_1, x_2)}{\partial x_2} = 0, \quad \frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} > 0, \quad \frac{\partial^2 f(x_1, x_2)}{\partial x_2^2} > 0$$

Is this enough? $\frac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_2}$ must also be taken into account.

Are these functions positive for all x_1, x_2 ? 1) $x_1^2 + x_1x_2 + x_2^2$, 2) $x_1^2 + 2x_1x_2 + x_2^2$, 3) $x_1^2 + 3x_1x_2 + x_2^2$

Positive Definite Matrix

- ▶ We can rearrange $ax_1^2 + 2bx_1x_2 + cx_2^2$ in the following manner,

$$f(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2 = a \left(x_1 + \frac{b}{a}x_2 \right)^2 + \left(c - \frac{b^2}{a} \right) x_2^2$$

$f(\bullet) > 0, \forall x_1, x_2 \neq 0$ when,

$$a > 0 \quad \text{and} \quad c - \frac{b^2}{a} > 0 \implies ac > b^2$$

$$\frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} > 0 \quad \text{and} \quad \frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} \frac{\partial^2 f(x_1, x_2)}{\partial x_2^2} > \left(\frac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_2} \right)^2$$

Positive Definite Matrix

- We can rearrange $ax_1^2 + 2bx_1x_2 + cx_2^2$ in the following manner,

$$f(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2 = a \left(x_1 + \frac{b}{a}x_2 \right)^2 + \left(c - \frac{b^2}{a} \right) x_2^2$$

$f(\bullet) > 0, \forall x_1, x_2 \neq 0$ when,

$$a > 0 \quad \text{and} \quad c - \frac{b^2}{a} > 0 \implies ac > b^2$$

$$\frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} > 0 \quad \text{and} \quad \frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} \frac{\partial^2 f(x_1, x_2)}{\partial x_2^2} > \left(\frac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_2} \right)^2$$

Verify this on the following functions: 1) $x_1^2 + x_1x_2 + x_2^2$, 2) $x_1^2 + 2x_1x_2 + x_2^2$, 3) $x_1^2 + 3x_1x_2 + x_2^2$

Positive Definite Matrix

$f(\bullet)$ can be expressed as $\mathbf{x}^T \mathbf{A} \mathbf{x}$, where $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, and $\mathbf{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$. $f(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$, when \mathbf{A} is positive definite.

$\mathbf{x}^T \mathbf{A} \mathbf{x}$ is called a *quadratic form*. For a symmetric matrix \mathbf{A} ,

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} = \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{i=1}^n \sum_{j=1, j \neq i}^n a_{ij} x_i x_j$$

In general, $\mathbf{A} \in \mathbb{R}^{n \times n}$ is positive definite, if:

- ▶ The eigenvalues of \mathbf{A} are all positive.
- ▶ The pivots (without row exchange) are all positive.

Positive Definite Matrix

$f(\bullet)$ can be expressed as $\mathbf{x}^T \mathbf{A} \mathbf{x}$, where $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, and $\mathbf{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$. $f(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$, when \mathbf{A} is positive definite.

$\mathbf{x}^T \mathbf{A} \mathbf{x}$ is called a *quadratic form*. For a symmetric matrix \mathbf{A} ,

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} = \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{i=1}^n \sum_{j=1, j \neq i}^n a_{ij} x_i x_j$$

In general, $\mathbf{A} \in \mathbb{R}^{n \times n}$ is positive definite, if:

- ▶ The eigenvalues of \mathbf{A} are all positive.
- ▶ The pivots (without row exchange) are all positive.

Show that any \mathbf{A} is positive definite if the symmetric matrix $\mathbf{A} + \mathbf{A}^T$ is positive definite. Note: *This should explain why we have only been talking about symmetric matrices.*

Matrix Norm

- ▶ Since matrices also form vector spaces, we can talk about norms of matrices, which extend the idea of sizes and distances to spaces of matrices.
- ▶ If we think of matrices a set of mn scalars, then we can use the same approach as vectors,

$$\|\mathbf{A}\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{\frac{1}{2}}$$

This is called the *Frobenius norm*.

Matrix Norm

- ▶ There are other norms defined for matrices that are very useful from the point of the view of linear transformation.
- ▶ These are called *induced matrix norms*, that looks at how matrices map vectors from the range to domain spaces.
- ▶ Let $\mathbf{A} \in \mathbb{R}^{m \times n} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, and $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$

$$\|\mathbf{y}\| = \|\mathbf{Ax}\| \leq C \|\mathbf{x}\|, \quad \forall \mathbf{x} \in \mathbb{R}^n$$

C is the maximum factor by which \mathbf{A} amplifies the vector \mathbf{x} .

- ▶ The induced norm of a matrix is defined as,

$$\|\mathbf{A}\|_p = \max_{\mathbf{x}} \frac{\|\mathbf{Ax}\|_p}{\|\mathbf{x}\|_p} = \max_{\|\mathbf{x}\|_p=1} \|\mathbf{Ax}\|_p$$

Matrix Norm

Consider a matrix $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = [\mathbf{a}_1 \quad \mathbf{a}_2] = \begin{bmatrix} \tilde{\mathbf{a}}_1^T \\ \tilde{\mathbf{a}}_2^T \end{bmatrix}$.

$$\|\mathbf{A}\|_1 = \max_i \|\mathbf{a}_i\|_1$$

$$\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}}$$

$$\|\mathbf{A}\|_\infty = \max_i \|\tilde{\mathbf{a}}_i\|_1$$

