Polynomials Optimization via Linear Matrix Inequality

A. M. S. Ang

First created: 2014. Last update: 2017-Feb

My own notes from the course **Polynomials Optimization via Linear Matrix Inequality** I took in HKU in 2014.

1 Polynomial

1.1 Single variable polynomial

For single variable polynomial $f(x) : \mathbb{R} \to \mathbb{R}$, it has the general form

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + ... + a_n x^n = \sum_{i=0}^{n} a_i x^i$$

n is the degree of the polynomial with coefficients $a_0, a_1, a_2, ..., a_n$.

For a polynomial of degree n , the coefficient a_n cannot be zero, otherwise it is not a polynomial in degree n. For other coefficients a_i $(i \neq n)$, there is no such restriction.

When $n = 1$, $f(x) = a_0 + a_1 x$ is called $affine\ function$. If $a_0 = 0$, $f(x) = a_1 x$ is called $linear\ function$.
When $n = 2$, $f(x) = a_0 + a_1 x + a_2 x^2$ is called $quadratic\ function$.
When $n = 3$, $f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$ is called $cubic\ function$.
When $n = 4$, $f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$ is called $quartic\ function$.

The solutions of the equation

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 + ... + a_n x^n = 0$$

are the *roots* of this polynomial equations.
When $n = 1$, $f(x) = a_0 + a_1 x = 0$ has the root $x = -\dfrac{a_0}{a_1}$
When $n = 2$, $f(x) = a_0 + a_1 x + a_2 x^2$ has the roots $x = \dfrac{-a_1 \pm \sqrt{a_1^2 - 4a_3 a_0}}{2a_2}$
When $n = 3$ and $n = 4$, the equations can be solved analytically , but the solution is very complicated.
When $n \geq 5$, there is no analytical solution (Abel's impossibility theorem).

1.2 Multivariable Polynomial

A $m-$degree n-variable polynomial is a mapping $\mathbb{R}^n \to \mathbb{R}$. Notice that now x is a $n-$dimensional vector $x = [x_1\, x_2\, ...\, x_n]$
e.g. $x_2 - \sqrt{4} x_1^2 + x_3^3$

The number of monomials in a polynomial in $n-$variables of degree m : $\sigma(n, m) = C_n^{n+m} = \dfrac{(n+m)!}{n!m!}$
e.g. 1-variable degree 4 polynomial $\sigma(1, 4) = \dfrac{5!}{1!4!} = 5 : 1, x_1, x_1^2, x_1^3, x_1^4$
e.g. 2-variable degree 2 polynomial $\sigma(2, 2) = \dfrac{4!}{2!2!} = 6 : 1, x_1, x_2, x_1^2, x_1 x_2, x_2^2$

1.3 Positive Definite Polynomial and Polynomial Optimization Problem

A polynomial is *positive definite* if $p(x) > 0\ \forall x$. That means the $p(x)$ is positive for any input x.

A polynomial is a sum of squares of polynomials (SOS) if there exists polynomials $p_1(x), p_2(x), ...$ that

$$p(x) = \sum p_i(x)^2 > 0$$

For a polynomial $p(x)$ is SOS, it must be semi-positive

The polynomial optimization problem is to find the optimizer of a polynomial

$$\text{find } x \text{ that maximize/minimize } p(x)$$

For example, find the minimum value of a SOS polynomial

$$\text{find inf} p(x) \text{ for } p(x) = \sum p_i(x)^2$$

1.4 Vector Representation of polynomial

For a polynomial $p(x)\,:\,\mathbb{R}^n \rightarrow \mathbb{R}.$, it can be expressed as the product between coefficient vector and basis vector

$$p(x) = c^T x$$

e.g. $x_2 - \sqrt{4}x_1^2 + x_3^3 = \begin{bmatrix} 1 & -\sqrt{4} & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ x_1^2 \\ x_3^3 \end{bmatrix}$

Vector representation of polynomial is not very useful in polynomial optimization. We cannot determine positivity of the polynomial from vector representation, so the *matrix representation* is considered.

2 Square Matrix Representation of Polynomials

For a polynomial $p(x)$, it can be expressed as

$$p(x) = \left(x^{\langle m\rangle}\right)^T \left[P + L\left(\alpha\right)\right] x^{\langle m\rangle}$$

$$P \in \mathbb{R}^{\sigma(n,m)} \text{ and } L \in \mathcal{L} = \left\{ L \in \mathbb{S}^{\sigma(n,m)} \middle| \left(x^{\langle m\rangle}\right)^T L x^{\langle m\rangle} = 0 \right\}$$

2.1 Power Vector $x^{\langle m\rangle}$

Let n denote the number of variables and m denote the degree.

For a degree m polynomial, the power vector $x^{\langle m\rangle}$ is the vector containing all *monomials* in x with degree less than or equal to m . *Lexicographical ordering* is used :

For $n = 1$, $m = 0$, $x^{\langle m\rangle} = \left[1\right]$

For $n = 1$, $m = 1$, $x^{\langle m\rangle} = \left[1\, x_1\right]$

For $n = 2$, $m = 1$, $x^{\langle m\rangle} = \left[1\, x_1\, x_2\right]$

For $n = 3$, $m = 1$, $x^{\langle m\rangle} = \left[1\, x_1\, x_2\, x_3\right]$

For $n = 1$, $m = 2$, $x^{\langle m\rangle} = \left[1\, x_1\, x_1^2\right]$

For $n = 2$, $m = 2$, $x^{\langle m\rangle} = \left[1\, x_1\, x_2\, x_1^2\, x_1x_2\, x_2^2\right]$

For $n = 3$, $m = 2$, $x^{\langle m\rangle} = \left[1\, x_1\, x_2\, x_3\, x_1^2\, x_1x_2\, x_1x_3\, x_2^2\, x_2x_3\, x_3^2\right]$

For $n = 1$, $m = 3$, $x^{\langle m\rangle} = \left[1\, x_1\, x_1^2\, x_1^3\right]$

For $n = 2$, $m = 3$, $x^{\langle m\rangle} = \left[1\, x_1\, x_2\, x_1^2\, x_1x_2\, x_2^2\, x_1^3\, x_1^2x_2\, x_1x_2^2\, x_2^3\right]$

For $n = 3$, $m = 3$, $x^{\langle m\rangle} = \left[1\, x_1\, x_2\, x_3\, x_1^2\, x_1x_2\, x_1x_3\, x_2^2\, x_2x_3\, x_3^2\, x_1^3\, x_1^2x_2\, x_1^2x_3\, x_1x_2^2\, x_1x_2x_3\, x_1x_3^2\, x_2^3\, x_2^2x_3\, x_2x_3^2\, x_3^3\right]$

2.2 Quadratic forms / Square matrix representation of polynomial

A polynomial $p(x) : \mathbb{R}^n \to \mathbb{R}$ can be expressed as the product of basis power vector to a matrix P. The matrix P is a square matrix :

$$p(x) = \left(x^{\langle m\rangle}\right)^T P x^{\langle m\rangle} = \sum_{i\,,j=1}^{n} P_{ij} x_i^{\langle m\rangle} x_j^{\langle m\rangle}$$

EXAMPLE. $p(x) = 1 + x_1 + x_1^2$

$$p(x) = \begin{bmatrix} 1 & x_1 \end{bmatrix} \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \end{bmatrix}$$

EXAMPLE. $p(x) = x_1 + x_2$

$$p(x) = \begin{bmatrix} 1 & x_1 & x_2 \end{bmatrix} \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0 \\ 0.5 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix}$$

EXAMPLE. $p(x) = 7 - 3x_2^2 + 9x_1^3$

$$p(x) = \begin{bmatrix} 1 & x_1 & x_2 & x_1^2 \end{bmatrix} \begin{bmatrix} 7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4.5 \\ 0 & 0 & -3 & 0 \\ 4.5 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \end{bmatrix}$$

Notice that the representation of a polynomial is *not unique* , for example

$$p(x) = 1 + x_1 + x_1^2 = \begin{bmatrix} 1 & x_1 \end{bmatrix} P \begin{bmatrix} 1 \\ x_1 \end{bmatrix} \quad P = \left\{ \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right\}$$

Because symmetric matrix has many useful properties, so symmetric matrix is used to present the polynomial.

Therefore, for a polynomial $p(x)$, it can be expressed as the quadratic forms (product of basis power vector $x^{\langle m\rangle}$ and matrix P where $P \in \mathbb{S}^{\sigma(n,m)}$) .

The *complete representation* of a polynomial consists of 2 symmetric matrices P and L . L is a matrix such that $x^{\langle m\rangle}$ is inside its kernel :

$$\left(x^{\langle m\rangle}\right)^T L x^{\langle m\rangle} = 0$$

That means $x^{\langle m\rangle} \in \ker L$.

Therefore the polynomial $p(x)$ can be represented as

$$p(x) = \left(x^{\langle m\rangle}\right)^T \Big[P + L\Big] x^{\langle m\rangle}$$

Since there are many possible L , all possible L form a set $\mathcal{L}$

$$L \in \mathcal{L} = \left\{ L \in \mathbb{S}^{\sigma(n,m)} \middle| \left(x^{\langle m\rangle}\right)^T L x^{\langle m\rangle} = 0 \right\}$$

The matrix L can be found by direct computation.

EXAMPLE. $p(x) = 1 + x_1 + x_2^3$

$$p(x) = 1 + x_1 + x_1^3 = \begin{bmatrix} 1 & x_1\, x_1^2 \end{bmatrix} \begin{bmatrix} 1 & 0.5 & 0 \\ 0.5 & 0 & 0.5 \\ 0 & 0.5 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_1^2 \end{bmatrix}$$

Since $L \in \mathbb{S}$ so we can let $L = \begin{bmatrix} l_{11} & l_{12} & l_{13} \\ l_{12} & l_{22} & l_{23} \\ l_{13} & l_{23} & l_{33} \end{bmatrix}$ and thus $\left(x^{\langle m \rangle}\right)^T L x^{\langle m \rangle}$ is

$$\begin{bmatrix} 1 & x_1 \, x_1^2 \end{bmatrix} \begin{bmatrix} l_{11} & l_{12} & l_{13} \\ l_{12} & l_{22} & l_{23} \\ l_{13} & l_{23} & l_{33} \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_1^2 \end{bmatrix} = \begin{bmatrix} 1 & x_1 \, x_1^2 \end{bmatrix} \begin{bmatrix} l_{11} + l_{12}x_1 + l_{13}x_1^2 \\ l_{12} + l_{22}x_1 + l_{23}x_1^2 \\ l_{13} + l_{23}x_1 + l_{33}x_1^2 \end{bmatrix} = \begin{matrix} l_{11} \\ +2l_{12}x_1 \\ +\left(2l_{13} + l_{22}\right)x_1^2 \\ +2l_{23}x_1^3 \\ +l_{33}x_1^4 \end{matrix}$$

Since the result has to be zero, thus $\begin{cases} l_{11} = l_{12} = l_{23} = l_{33} = 0 \\ 2l_{13} + l_{22} = 0 \end{cases}$ and $L = \begin{bmatrix} l_{11} & l_{12} & l_{13} \\ l_{12} & l_{22} & l_{23} \\ l_{13} & l_{23} & l_{33} \end{bmatrix} = \begin{bmatrix} 0 & 0 & l_{13} \\ 0 & -2l_{13} & 0 \\ l_{13} & 0 & 0 \end{bmatrix}$ and l_{13} can be any number. Denote l_{13} as α :

$$p(x) = 1 + x_1 + x_1^3 = \begin{bmatrix} 1 & x_1 \, x_1^2 \end{bmatrix} \left(\begin{bmatrix} 1 & 0.5 & 0 \\ 0.5 & 0 & 0.5 \\ 0 & 0.5 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & \alpha \\ 0 & -2\alpha & 0 \\ \alpha & 0 & 0 \end{bmatrix} \right) \begin{bmatrix} 1 \\ x_1 \\ x_1^2 \end{bmatrix}$$

3 Incomplete Basis Representation : Form and Newton's Polytope

3.1 Motivation of incomplete basis representation

Consider 3-variables degree 6 polynomial $p(x) = x_1^6 + x_2^6 + x_3^6$, then the power vector is

$$x^{\langle m \rangle} = \left[1 \, x_1 \, x_2 \, x_3 \, x_1^2 \, x_1x_2 \, x_1x_3 \, x_2^2 \, x_2x_3 \, x_3^2 \, x_1^3 \, x_1^2x_2 \, x_1^2x_3 \, x_1x_2^2 \, x_1x_2x_3 \, x_1x_3^2 \, x_2^3 \, x_2^2x_3 \, x_3^3\right] \in \mathbb{R}^{20}$$

It has $\dfrac{6!}{3!3!} = 20$ dimension. Then the P and L matrix will be 20×20 , which is VERY BIG.

But it can be observed that, there will be many zeros in the P and L matrix. For example, since the polynomial does not have constant term, thus $p_{11} = l_{11} = 0$.

This suggest that, in the power vector, the "1" is actually redundant and can be removed. Therefore this lead to *incomplete basis*.

3.2 Incomplete Basis Representation of polynomial by form / homogenous basis

A *form* is a polynomial that all the terms are having the same degree.
For example, $x_1^2 + x_1x_2 + x_2^2$, $x_2^3 - 5x_1x_3^2 + 0.2x_1x_2x_3$ are forms.

Forms also called as *homogenous plynomial*.

The degree of a homogenous polynomial can be determined by $x_i \leftarrow \lambda x_i$.
For example $x_1^2 + x_1x_2 + x_2^2$ is in degree 2 since $(\lambda x_1)^2 + (\lambda x_1)(\lambda x_2) + (\lambda x_2)^2 = \lambda^2 \left(x_1^2 + x_1x_2 + x_2^2\right)$

Then, for representing a polynomial up to degree $2m$, a incomplete power vector with homogenous monomial up to degree m can be used.
For example, $p(x) = x_1^6 + x_2^6 + x_3^6$, then the homogenous power vector is $x_{HOM}^{\langle m \rangle} = \left[x_1^3 \, x_2^3 \, x_3^3\right]$

$$p(x) = \begin{bmatrix} x_1^3 \\ x_2^3 \\ x_3^3 \end{bmatrix}^T \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + L \right\} \begin{bmatrix} x_1^3 \\ x_2^3 \\ x_3^3 \end{bmatrix}$$

In this case, the dimension of the basis vector reduced from 20 to 3. And the size of the P and L matrix are now 3×3.

Beside the homogenous vector, another way is to use the $Newton's\ Polytope$. Newton's Polytype is a tool for understanding and processing polynomials $p(x)\ :\ \mathbb{R}^n \to \mathbb{R}$.

3.3 Convex Hull

A *convex hull* of a set of points $V = \{v_1, v_2, ..., v_n\}$ is a set S of *all convex combinations* of its points. The mathematical definitions of convex hull of a set of points V is

$$\text{ConvexHull}\,(V) = conv\,(v_1, v_2, ..., v_n) = \left\{ \sum_{i=1}^{n} b_i v_i \middle| (\forall i\ :\ b_i \geq 0) \wedge \sum_{i=1}^{n} b_i = 1 \right\}$$

Explanations
The vertices v_i are *vectors*. e.g. for 2-dimensional case $v_i = [x_i\, y_i]$
$\sum b_i v_i$ is a linear combination of the vector v
$\sum b_i = 1$ meansthat the combination is "convex".

3.4 Newton's Polytope

Consider a polynomial $p(x)\ :\mathbb{R}^n \to \mathbb{R}$, it can be represented by a square matrix $p(x) = \left(x^{\langle m \rangle}\right)^T \Big[P{+}L\,(\alpha)\Big] x^{\langle m \rangle}$

This polynomial $p(x)$ is non-negative definite iff it is *sum of squares* (SOS) : $p(x) = \sum_{i=1}^{l} \tilde{p}_i(x)^2$. That means, there exist some polynomials $\tilde{p}_i(x)$, that the original polynomial $p(x)$ can be expressed as the sum of these polynomials. Since squared-polynomial cannot be negative, thus if $p(x)$ is SOS, then $p(x) \geq 0$.

A polynomial can also be represented as a sums of power vectors. This is true for all both polynomials $p(x)$ and $\tilde{p}(x)$. So $p(x) = \sum_{j=1}^{N} c_j x^{k(j)}$ and $\tilde{p}_i(x) = \sum_{j=1}^{N} \tilde{c}_j x^{\tilde{k}(j)}$

Since $p(x)$ is SOS $p(x) = \sum_{i=1}^{l} \tilde{p}_i(x)^2$. Therefore there should be some relationship between the power vectors of p and $\tilde{p}_i$.

It can be shown that the relationship between the powers of the basis vector x for p and $\tilde{p}_i$ are related by the Newton's Polytope as follows :

$$2\tilde{k}(j) \in conv\Big[k(1), ..., k(N)\Big]$$

That means, to express a polynomial as SOS , the new basis vectors are selected such that

"the 2 times of their power indices are inside the convex hull generated by
the power indices of the basis vectors of the original polynomial".

EXAMPLE
Consider a polynomial $p(x) = x_1^2 + x_1^3 + x_1^6 + x_2^4$
That is equal to

$$p(x) = x_1^2 + x_1^3 + x_1^6 + x_2^4 = x_1^2 x_2^0 + x_1^3 x_2^0 + x_1^6 x_2^0 + x_1^0 x_2^4$$

Expressed as sum of power vectors

$$p(x) = x^{\begin{bmatrix} 2 \\ 0 \end{bmatrix}} + x^{\begin{bmatrix} 3 \\ 0 \end{bmatrix}} + x^{\begin{bmatrix} 6 \\ 0 \end{bmatrix}} + x^{\begin{bmatrix} 0 \\ 4 \end{bmatrix}}$$

Therefore the power indices are

$$\left\{ \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 6 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \end{bmatrix} \right\}$$

And the Newton's Polytope is

$$\text{conv}\left\{\begin{bmatrix}2\\0\end{bmatrix},\begin{bmatrix}3\\0\end{bmatrix},\begin{bmatrix}6\\0\end{bmatrix},\begin{bmatrix}0\\4\end{bmatrix}\right\}$$

Since $p(x)$ is a degree 6 polynomial with 2 variables. Therefore the basis power vector of this polynomial is

$$x = \left[1\ x_1\ x_2\ x_1^2\ x_1x_2\ x_2^2\ x_1^3\ x_1^2x_2\ x_1x_2^2\ x_2^3\right]$$

The power indices are

$$\left\{\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix},\begin{bmatrix}2\\0\end{bmatrix},\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}0\\2\end{bmatrix},\begin{bmatrix}3\\0\end{bmatrix},\begin{bmatrix}2\\1\end{bmatrix},\begin{bmatrix}1\\2\end{bmatrix},\begin{bmatrix}0\\3\end{bmatrix}\right\}$$

Not all components are needed to represent the polynomial. To check which components of the vectors should be included, solve

$$2\left\{\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix},\begin{bmatrix}2\\0\end{bmatrix},\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}0\\2\end{bmatrix},\begin{bmatrix}3\\0\end{bmatrix},\begin{bmatrix}2\\1\end{bmatrix},\begin{bmatrix}1\\2\end{bmatrix},\begin{bmatrix}0\\3\end{bmatrix}\right\}\overset{?}{\in}\text{conv}\left\{\begin{bmatrix}2\\0\end{bmatrix},\begin{bmatrix}3\\0\end{bmatrix},\begin{bmatrix}6\\0\end{bmatrix},\begin{bmatrix}0\\4\end{bmatrix}\right\}$$

Note that there is a '2' on the left hand side. After solving

$$\left\{\begin{bmatrix}2\\0\end{bmatrix},\begin{bmatrix}4\\0\end{bmatrix},\begin{bmatrix}2\\2\end{bmatrix},\begin{bmatrix}0\\4\end{bmatrix},\begin{bmatrix}6\\0\end{bmatrix}\right\}\in\text{conv}\left\{\begin{bmatrix}2\\0\end{bmatrix},\begin{bmatrix}3\\0\end{bmatrix},\begin{bmatrix}6\\0\end{bmatrix},\begin{bmatrix}0\\4\end{bmatrix}\right\}$$

Therefore only the following components are required

$$b(x) = \left[x_1\ x_1^2\ x_1x_2\ x_2^2\ x_1^3\right]$$

3.5 Application of Newton's Polytope on Matrix Polynomial representation

Consider the polynomial $p(x) = x_1^2 + x_1^3 + x_1^6 + x_2^4$. In matrix representation

$$p(x) = \left(x^{\langle m\rangle}\right)^T P x^{\langle m\rangle}, \text{ where } x^{\langle m\rangle} = \left[1\ x_1\ x_2\ x_1^2\ x_1x_2\ x_2^2\ x_1^3\ x_1^2x_2\ x_1x_2^2\ x_2^3\right]$$

the matrix P is 10×10.

By applying Newton's Polytope , the basis power vector is changed form $x^{\langle m\rangle}$ to $b(x) = \left[x_1\ x_1^2\ x_1x_2\ x_2^2\ x_1^3\right]$, therefore

$$p(x) = (b(x))^T \tilde{P} b(x)$$

the matrix $\tilde{P}$ is now 5×5 , which is much smaller.

So the application of Newton's Polytope is to reduce the dimensions. For a complete representation of polynomial $p(x)$, the matrix representation is $p(x) = \left(x^{\langle m\rangle}\right)^T \left[P + L(\alpha)\right] x^{\langle m\rangle}$. By using Newton's Polytope, the new representation $p(x) = (b(x))^T \left[\tilde{P} + \tilde{L}(\alpha)\right] b(x)$, the size of $\tilde{P}$ and $\tilde{L}$ will be smaller, and the dimension of the free parameter vector α will also be greatly reduced.

4 Matrix Inequality

4.1 Positive Definite Matrix

The inequality $\neq$ basically means $>$ or $<$. For scalar case, it can be $a > b$ or $a < b$. But the inequality for matrix is not that simple.

First, we cannot compare a matrix with a scalar directly. For "$A > 0$ " (where 0 is scalar), it is technically incorrect.

We can compare a matrix with another matrix. So first consider comparing a matrix to a *zero matrix* $\mathbf{0}$.

For a $m \times n$ matrix A, $A > \mathbf{0}$ means the matrix A is *positive definite*. The definition of positive definite is : $A > 0$ if if $x^T Ax$ is positive for all non-zero column vector x

$$A > \mathbf{0} \iff x^T Ax > 0$$

Similarly, matrix A is *positive semidefinite* if the $>$ sign is changed into $\geq$ sign.

$$A \geq \mathbf{0} \iff x^T Ax \geq 0$$

Using the similar definitions , *negative definite* and *negative semidefinite* can be defined :

$$A < \mathbf{0} \iff x^T Ax < 0$$

$$A \leq \mathbf{0} \iff x^T Ax \leq 0$$

For scalar case, we have the *Law of Trichotomy* (for real numbers)

$$\text{For } a > b\,,\ a = b\,,\ a < b \text{ only one of it holds}$$

Therefore, $a \ngtr b$ also means $a < b$ and $a \nless b$ also means $a \geq b$.

But that is not true for matrix.

$$A \nless \mathbf{0} \implies A > \mathbf{0} \text{ is not ture}$$

Indeed, for matrix inequality, a matrix A can also be *indefinite*

$$A \text{ is indefinite if } A \ngtr \mathbf{0} \text{ AND } A \nless \mathbf{0}$$

Therefore, a matrix A can be

$$A \text{ can be} \begin{cases} \text{positive definite} & A > \mathbf{0} & x^T Ax > 0 & \forall x \neq 0 \\ \text{positive semidefinite} & A \geq \mathbf{0} & x^T Ax \geq 0 & \forall x \\ \text{negative definite} & A < \mathbf{0} & x^T Ax < 0 & \forall x \neq 0 \\ \text{negative semidefinite} & A \leq \mathbf{0} & x^T Ax \leq 0 & \forall x \\ \text{indefinite} & A \ngtr \mathbf{0} \text{ AND } A \nless \mathbf{0} & & \forall x \end{cases}$$

4.2 Comparing two matrix

The comparison of two matrix A and B is equivalent to the comparison between $A - B$ and $\mathbf{0}$

4.3 How to check a matrix is positive definite or not

There are many ways to check a matrix A is positive definite or not.

Method. 1 By definition. The first way is to use the definition $A > \mathbf{0} \iff x^T Ax > 0$

[e.g.1] $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

For any vector $x = \begin{bmatrix} a \\ b \end{bmatrix}$

$x^T Ax = [a\ b] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$

$x^T Ax = [a\ b] \begin{bmatrix} a \\ b \end{bmatrix}$

$x^T A x = a^2 + b^2$
Since x can not be zero vector (by the definition), so the sum of squares $a^2 + b^2$ must be > 0
Thus matrix $A > \mathbf{0}$ and it is positive definite.

[e.g.2] $B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$
For any vector $x = \begin{bmatrix} a \\ b \end{bmatrix}$,
$x^T B x = [a\ b] \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$
$x^T B x = [a - b\ \ -a + b] \begin{bmatrix} a \\ b \end{bmatrix}$
$x^T B x = a^2 - ab - ab + b^2 = (a - b)^2 > 0$
Thus matrix $B > \mathbf{0}$ and it is positive definite.

[e.g.3] $C = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$
For any vector $x = \begin{bmatrix} a \\ b \end{bmatrix}$, $x^T B x = [a\ b] \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = [a + 2b\ 2a + b] \begin{bmatrix} a \\ b \end{bmatrix} = a^2 + 4ab + b^2 = (a + b)^2 + 2ab$
Although $(a + b)^2 > 0$ but $2ab$ can be negative.
Indeed C is not positive definite. For $x = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $x^T A x = (1 + (-1))^2 + 2(1)(-1) = -2 < 0$
Thus matrix $C \not> \mathbf{0}$ and it is not positive definite.
In this stage, we just know that C is not positive definite, we don't know is C negative definite or not, need to check.

[e.g.4] $D = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$
For any vector $x = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, $x^T D x = [a\ b\ c] \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = [a\ b\ c] \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$
$= [a\ b\ c] \begin{bmatrix} 2a - b \\ -a + 2b - c \\ -b + 2c \end{bmatrix} = 2a^2 - ab - ab + 2b^2 - bc - bc + 2c^2$
$= 2a^2 + 2b^2 + 2c^2 + -2ab - 2bc$
$= a^2 + c^2 + (a^2 - 2ab + b^2) + (b^2 - 2bc + c^2)$
$= a^2 + c^2 + (a - b)^2 + (b + c)^2 > 0$ (it is a sum of squares)
Thus matrix $D > \mathbf{0}$ and it is positive definite.

Method 2. All principal minor $>$ 0. Having to do computation of a 3×3 matrix may be tedious. So the second way to determine is a matrix positive definite is to use the *principal minor*.

For a square matrix $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$, the principal minors are $\mu_1 = [a_{11}]$, $\mu_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$,
$\mu_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, ... $\mu_n = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$

By using principal minors, a matrix is positive definite if determinant of allprincipal minors are positive : $A > \mathbf{0} \iff \det \mu_i > 0\ \forall i$

For matrix $D = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$

$\det \mu_1 = 2 > 0$, $\det \mu_2 = \det \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = 5 > 0$, $\det \mu_3 = \det \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = 4 > 0$. All $\det\mu_i > 0$ so $D > \mathbf{0}$

Method 3. All eigenvalues > 0. Eigenvalues help to determine a matrix is positive definite or not. A matrix is positive definite if all eigenvalues are positive : $A > \mathbf{0} \iff \lambda_i(A) > 0 \; \forall i$

For matrix $D = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$, the eigenvalues are 0.5858 , 2 , 3.4142 . All $\lambda > 0$. So $D > \mathbf{0}$

Similarly, for negative definite matrix, all eigenvalues will be negative.

For indefinite matrix, some eigenvalues will be positive and some will be negative. So $\exists i, j : \lambda_i(A)\lambda_j(A) < 0$

Method 4. Gram matrix / Symmetric dyads. If matrix A can be factorized as product of a symmetric matrix : $A = B^T B$, then A is positive definite since $x^T A x = x^T B^T B x = ||Bx||_2^2 \geq 0$. In this case, the matrix A is a *Gram matrix*.

4.4 Summary

$$A > \mathbf{0} \begin{cases} x^T A x > 0 & \forall x \neq 0 \\ \lambda_i(A) > 0 & \forall i \\ \det \mu_i > 0 & \forall i \end{cases} \qquad A \geq \mathbf{0} \begin{cases} x^T A x \geq 0 & \forall x \\ \lambda_i(A) \geq 0 & \forall i \\ \det \mu_i \geq 0 \text{ and } a_{ii} \geq 0 & \forall i \end{cases}$$

$$A < \mathbf{0} \begin{cases} x^T A x < 0 & \forall x \neq 0 \\ \lambda_i(A) < 0 & \forall i \\ \det \mu_i < 0 & \forall i \end{cases} \qquad A \leq \mathbf{0} \begin{cases} x^T A x \leq 0 & \forall x \\ \lambda_i(A) \leq 0 & \forall i \\ \det \mu_i \leq 0 \text{ and } a_{ii} \leq 0 & \forall i \end{cases}$$

$$A \text{ indefinite : } \begin{cases} A \nleq \mathbf{0} \text{ AND } A \ngeq \mathbf{0} \\ \exists i, j : \lambda_i(A)\lambda_i(A) < 0 \end{cases}$$

5 Polynomials Optimization via Linear Matrix Inequality

5.1 The problem

For a $n-$variable degree m polynomial $f(x) : \mathbb{R}^n \to \mathbb{R}$, the problem is to find its optimal point , for example, minimize the polynomial

$$\min p(x)$$

It is the same as the maximization of the negative of the polynomial :

$$\max -p(x)$$

If $p(x)$ is a positive polynomial, then sometime it is interesting and meaningful to find the global minimum

$$\min p(x)$$

Notice that finding the minimum of a positive polynomial is equivalent to find the maximum t such that

$$p(x) - t \geq 0$$

5.2 The LMI method

Recalled that a polynomial $p(x)$ can be represented in matrix form

$$p(x) = \left(x^{\langle m\rangle}\right)^T \left[P + L(\alpha)\right] x^{\langle m\rangle}$$

where $P = P^T$, $L = L^T$, and $\left(x^{\langle m\rangle}\right)^T L(\alpha)\, x^{\langle m\rangle} = 0$

Then since

$$\min p(x) \iff \max t \ \text{ s.t.} p(x) - t \geq 0$$

Therefore the new polynomial $p(x) - t$ can also be represented in matrix form

$$p(x) - t = \left(x^{\langle m\rangle}\right)^T \left[P + L(\alpha) - tO\right] x^{\langle m\rangle}$$

where

$$\left(x^{\langle m\rangle}\right)^T O x^{\langle m\rangle} = 1 \quad\iff O = \begin{bmatrix} 1 & 0 \\ 0 & O_{22} \end{bmatrix}, O_{22} \text{ is zero matrix}$$

Then in that case

$$p(x) - t \geq 0 \iff P + L(\alpha) - tO \text{ is positive semidefinite}$$

Therefore, to find the minimum value of $p(x)$

$$\text{find } t{,}\alpha \ s.t. P + L(\alpha) - tO \geq 0$$

and the problem is to find the maximum t

In the same sense as minimization, the maximization problem can be treated as

$$\max p(x) \iff \min t \ \text{ s.t.} t - p(x) \geq 0$$

Therefore,

$$t - p(x) = \left(x^{\langle m\rangle}\right)^T \left[tO - P + L(\alpha)\right] x^{\langle m\rangle}$$

Then in that case

$$t - p(x) \geq 0 \iff tO - P + L(\alpha) \ \text{ is positive semidefinite}$$

Therefore, to find the maximum value of $p(x)$

$$\text{find } t{,}\alpha \ s.t. tO - P + L(\alpha) \geq 0$$

and the problem is to find the minimum t