Quadratic form:-

linear form:-

$$ax + by + cz$$

The way we will convert this equation in matrix:-

$$V \cdot X^T$$

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \begin{bmatrix} x & y & z \end{bmatrix} \rightarrow ax + by + cz.$$

quadratic form:-

$$ax^2 + 2bxy + cy^2$$

In matrix form

$$X^T M X$$

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \rightarrow$$

$$\Rightarrow \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} ax + by \\ bx + cy \end{bmatrix} \rightarrow ax^2 + bxy + bxy + cy^2$$

M can be a huge dimensional matrix. So this form $X^T M X$ allows us to represent a very complex quadratic function into compact form.

Positive semi definite matrix (PSD):-

Every square matrix M encodes to a quadratic function.

$$x \mapsto x^T M x = \sum M_{ij} x_i x_j$$

Here,

$$M \in \mathbb{R}^{d \times d} \qquad x \in \mathbb{R}^d$$

A symmetric matrix M is positive semi-definite if

$$x^T M x \geq 0 \quad , \forall x$$

in other words if the quadratic form of that matrix is greater than or equal zero

<u>For example</u> :

$$M = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$x^T M x \rightarrow [x_1 \quad x_2] \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\Rightarrow [x_1 \quad x_2] \begin{bmatrix} x_1 + x_2 \\ x_1 + x_2 \end{bmatrix}$$

$$\Rightarrow x_1^2 + x_1 x_2 + x_1 x_2 + x_2^2$$

$$\Rightarrow (x_1 + x_2)^2$$

$\forall x \in \mathbb{R}^n$, M's quadratic form will always greater than zero. So M will be a PSD.

* A diagonal matrix is PSD if & only if all the diagonal elements ≥ 0.

$$M_{ii} \geq 0$$

* if M is PSD then cM will be PSD
* if $M, N \in \mathbb{R}^{n \times n}$ & are PSD then $M+N$ is PSD.

It is not possible to check for every x that $x^T M x \geq 0$. So another way to check PSD is:-
A matrix is PSD if & only if it can be written as

$$M = UU^T \text{ for some matrix } U$$

<u>Quick check:-</u>

say $U \in \mathbb{R}^{r \times d}$ and $M = UU^T$

→ M is square → yes.

$M = UU^T$ & $U \in \mathbb{R}^{r \times d}$.

$\therefore M \in \mathbb{R}^{r \times r}$

→ M is symmetry → yes

$$M = M^T.$$

$$\begin{aligned} UU^T &= (UU^T)^T \\ &= (U^T)^T U^T \\ &= UU^T \end{aligned}$$

$(AB)^T = B^T A^T$.

→ $x^T M x \geq 0$ → yes.

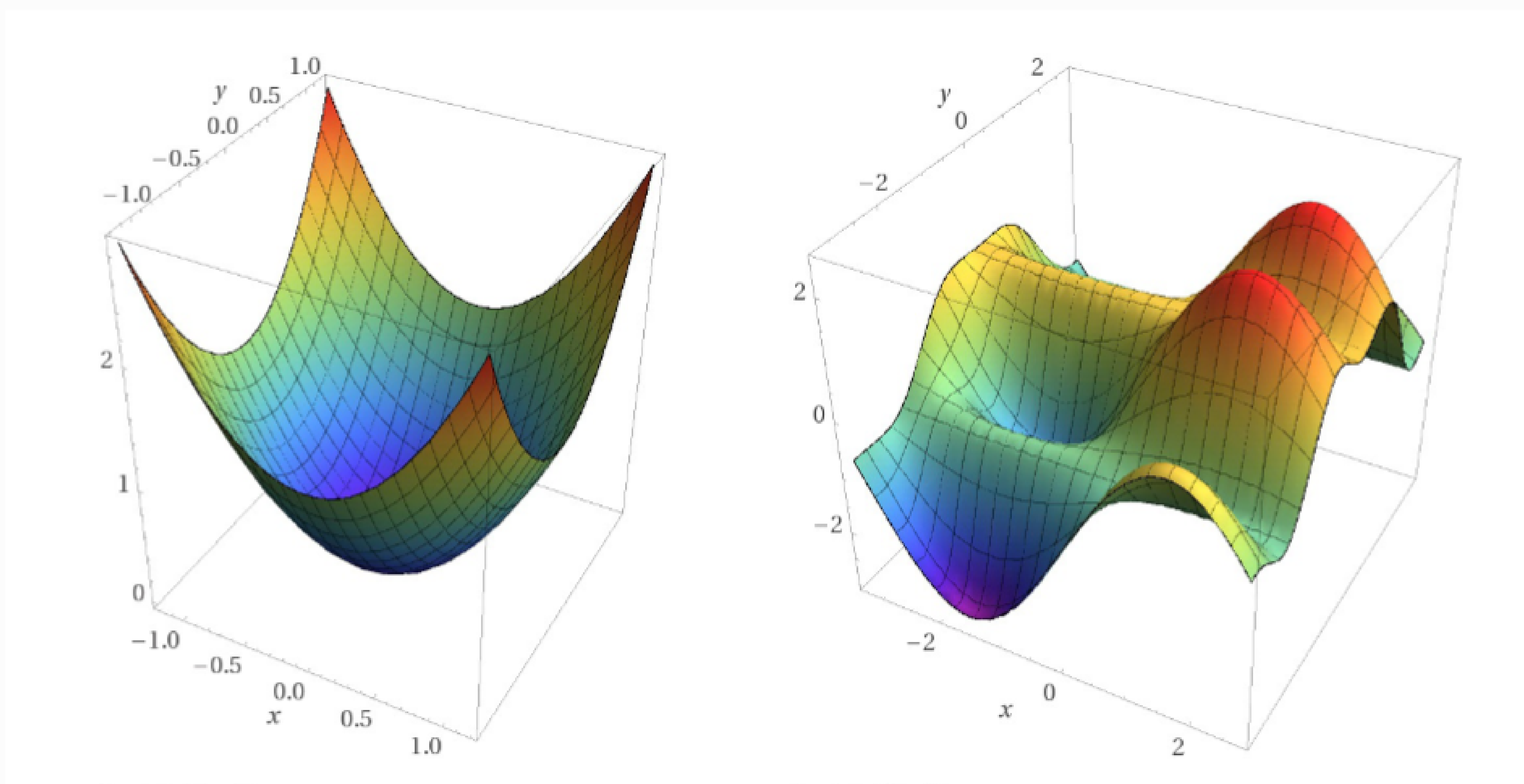
$$\begin{aligned} x^T M x &= x^T U U^T x \\ &= (x^T U)(U^T x) \end{aligned}$$

$A^T A = A^2$

$$= (U^T x)^T (U^T x)$$
$$= \|U^T x\|^2 \geq 0$$

Why PSD is important:-

It's very hard to optimize the non-convex function. PSD is a good indicator of convexity of a function.



convex non-convex

A function of several variables, $F(z)$ is convex if its second derivative matrix $H(z)$ is PSD.

$H(z)$ is also called Hessian matrix.

$$H(f(x_1, x_2, x_3 \ldots x_n)) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} & \cdots & \frac{\partial^2 f}{x_1 x_n} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \vdots \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n x_1} & \cdots & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$