c) a global maximizer for $f(\vec{x})$ if $Hf(\vec{x})$ is negative semidefinite on $\mathbb{R}^n$

d) a strict global maximizer for $f(\vec{x})$ if $Hf(\vec{x})$ is negative definite on $\mathbb{R}^n$.

How can we tell whether a matrix A is positive (negative) (semi) definite?

There are two ways:

- Using determinants
- Using eigenvectors

Method 1: Through determinants

Consider a two-dimensional case first. The n-dimensional case can be derived through induction.

Let $f_{xx} = a$, $f_{yy} = b$, $h = f_{xy} = f_{yx}$ then Hf would be positive definite if

$$x^T Hf x = (x_1, x_2)^T \begin{pmatrix} a & h \\ h & b \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} > 0$$

$\Rightarrow ax_1^2 + 2hx_1x_2 + bx_2^2 > 0$

Since we want to see whether itf is positive definite independently of $\vec{x}$, we can make $x_1 = tx_2$ for some $t \in \mathbb{R}$

So

$at^2x_2^2 + 2htx_2^2 + bx_2^2 > 0$

$(at^2 + 2ht + b)x_2^2 > 0$ which occurs for any $x_2 \neq 0$, and $(at^2 + 2ht + b) > 0$ $t \neq 0$

$\varphi(t) = at^2 + 2ht + b$

$\varphi'(t) = 2at + 2h$ $\qquad\qquad \varphi''(t) = 2a$

$a > 0 \quad \varphi(t^*)$ min

$a < 0 \quad \varphi(t^*)$ max

t^*

$\varphi'(t^*) = 2at^* + 2h = 0$

$t^* = -\dfrac{h}{a}$

$$\varphi(t^*) = a\left(-\frac{h}{a}\right)^2 + 2h\left(-\frac{h}{a}\right) + b$$

$$= \frac{h^2}{a} - \frac{2h^2}{a} + b$$

$$= \frac{-h^2}{a} + b = \frac{ab - h^2}{a}$$

We can conclude then that

$\varphi(t) > 0$ if $a > 0$ & $ab - h^2 > 0$

or

$$\boxed{a > 0 \quad \text{and} \quad \det\begin{pmatrix} a & h \\ h & b \end{pmatrix} > 0}$$

Condition taught in Calculus

Theorem: A 2×2 symmetric matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$$

is:

a) positive definite iff

$$a_{11} > 0 \quad \& \quad \det\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} > 0$$

b) negative definite iff

$$a_{11} < 0 \quad \& \quad \det\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} > 0$$

In the n-th dimensional case A is an $n \times n$ symmetric matrix. Define Δ_k to be the determinant of the upper left-hand corner $k \times k$ submatrix of A $1 \leq k \leq n$. The determinant Δ_k is called th kth principal minor of A

$$\begin{array}{c} \quad \Delta_1 \qquad \Delta_2 \qquad \Delta_3 \qquad\qquad\qquad \\ A = \left(\begin{array}{c|c|c|ccc} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ \hline a_{12} & a_{22} & a_{23} & \cdots & a_{2n} \\ \hline a_{13} & a_{23} & a_{33} & \cdots & a_{3n} \\ \hline \vdots & & & & \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{nn} \end{array}\right) \end{array}$$

$\Delta_1 = a_{11}$

$\Delta_2 = \det\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$

$\vdots$

$\Delta_n = \det A$

Theorem: If A is an $n \times n$-symmetric matrix and if Δ_k is the kth principal minor of A for $1 \leq k \leq n$, then:

a) A is positive definite iff $\Delta_k > 0$ for $k = 1, 2, \ldots, n$.

b) A is negative definite iff $(-1)^k \Delta_k > 0$ for $k = 1, 2, \ldots, n$ (i.e., the principal minors alternate in sign with $\Delta_1 < 0$).

Method 2: Through eigenvectors

Remember that an eigenvector of a matrix A is a vector $\vec{x}$ that satisfies

$$A\vec{x} = \lambda \vec{x} \qquad (1)$$

where λ is a scalar called the eigenvalue associated to $\vec{x}$, with $\vec{x} \neq \vec{0}$.

(1) can be rewritten as

$$(A - \lambda I)\vec{x} = \vec{0}$$

If we want a nontrivial solution, then

(2) $\det(A - \lambda I) = 0$ $\left\{ \begin{array}{l} \text{because the coefficients are} \\ \text{linearly dependent} \end{array} \right\}$

(2) will result in a n-th degree polynomial in λ, and thus there will be n roots $(\lambda_1, \ldots, \lambda_n)$.

Since there are an infinite number of vectors $\vec{x}$ that satisfy (1) for each λ_i, we normalize these vectors and refer to them as $\hat{x}_i$.

Example: Find the eigenvectors and eigenvalues of

$$A = \begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix}.$$

$$\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 2 \\ 2 & -1-\lambda \end{vmatrix} = (2-\lambda)(-1-\lambda) - 4 = 0$$

$$-2 - 2\lambda + \lambda + \lambda^2 - 4 = \lambda^2 - \lambda - 6 = 0$$

Since $\lambda^2 - \lambda - 6 = (\lambda - 3)(\lambda + 2) = 0$

$\lambda_1 = 3$, $\lambda_2 = -2$

Using $\lambda_1 = 3$:

$$\begin{pmatrix} 2-3 & 2 \\ 2 & -1-3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \underbrace{\begin{pmatrix} -1 & 2 \\ 2 & -4 \end{pmatrix}}_{\text{linearly dependent rows.}}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= \begin{pmatrix} -x_1 + 8x_2 \\ 2x_1 - 4x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow$$

$$-x_1 + 2x_2 = 2x_1 - 4x_2$$

$$-3x_1 = -6x_2$$

$$x_1 = 2x_2$$

Now, normalizing: $(x_1^2 + x_2^2 = 1)$

$$4x_2^2 + x_2^2 = 1 \Rightarrow x_2^2 = \frac{1}{5} \Rightarrow$$

$$x_2 = \frac{1}{\sqrt{5}} \text{ and } x_1 = \frac{2}{\sqrt{5}} \text{ or } \left(-\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right)$$

With $\lambda_2 = -2$ we get.

$$\hat{x}_2 = \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right) \text{ or } \hat{x}_2 = \left(\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}}\right)$$

Because $\hat{x}_i$ are unit vectors, $\hat{x}_i \cdot \hat{x}_i = 1$ and, additionally, $\hat{x}_i \cdot \hat{x}_j = 0$ $i \neq j$.

This is because if

$$A\hat{x}_i = \lambda_i \hat{x}_i \quad (a)$$
$$A\hat{x}_j = \lambda_j \hat{x}_j \quad (b)$$

pre-multiplying:

$$\hat{x}_j \cdot A\hat{x}_i = \hat{x}_j \cdot \lambda_i \hat{x}_i = \lambda_i \hat{x}_j \cdot \hat{x}_i = \lambda_i \hat{x}_i \cdot \hat{x}_j$$

$$\hat{x}_i \cdot A\hat{x}_j = \hat{x}_i \cdot \lambda_j \hat{x}_j = \lambda_j \hat{x}_i \cdot \hat{x}_j$$

Since $\hat{x}_j \cdot A\hat{x}_i$ and $\hat{x}_i \cdot A\hat{x}_j$ are scalars and A is symetric, then

$$\hat{x}_j \cdot A\hat{x}_i = \hat{x}_i \cdot A\hat{x}_j$$

$\Rightarrow$

$$\lambda_i x_i \cdot x_j = \lambda_j x_i \cdot x_j$$

$$\underbrace{(\lambda_i - \lambda_j)}_{\neq 0}\underbrace{(x_i \cdot x_j)}_{0} = 0$$

* Now the idea is to transform $\vec{x} \cdot A\vec{x}$ into $\lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots \lambda_n y_n^2$ so that everything depends on λ_i.

eigenvectors of A.

Let $B = [\hat{x}_1 \; \hat{x}_2 \; \dots \; \hat{x}_n]$, then set

$\vec{x} = B\vec{y}$, then

$$\vec{x} \cdot A\vec{x} = \vec{x}^T A \vec{x} = (B\vec{y})^T A (B\vec{y})$$

$$= \vec{y}^T B^T A B \vec{y}$$

Now, if $R = B^T A B$, what is the form of R?

$$R = \begin{bmatrix} \hat{x}_1^T \\ \hat{x}_2^T \\ \vdots \\ \hat{x}_n^T \end{bmatrix} A \begin{bmatrix} \hat{x}_1 & \hat{x}_2 & \cdots & \hat{x}_n \end{bmatrix}$$

$$= \begin{bmatrix} \hat{x}_1^T \\ \hat{x}_2^T \\ \vdots \end{bmatrix} \begin{bmatrix} A\hat{x}_1 & A\hat{x}_2 & \cdots & A\hat{x}_n \end{bmatrix}$$

$$= \begin{bmatrix} \hat{x}_1^T \\ \hat{x}_2^T \\ \vdots \end{bmatrix} \begin{bmatrix} \lambda_1 \hat{x}_1 & \lambda_2 \hat{x}_2 & \cdots & \lambda_n \hat{x}_n \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 \hat{x}_1^T \hat{x}_1 & \lambda_2 \hat{x}_1^T \hat{x}_2 & \cdots & \lambda_n \hat{x}_1^T \hat{x}_n \\ \lambda_1 \hat{x}_2^T \hat{x}_1 & \lambda_2 \hat{x}_2^T \hat{x}_2 & \cdots & \\ \vdots & & \ddots & \\ & & & \lambda_n \hat{x}_n^T \hat{x}_n \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & 0 & \cdots & 0 \\ \vdots & & & & \ddots & \lambda_n \end{bmatrix}$$

Therefore, we can conclude;

a) A is positive (negative) definite, iff every eigenvalue of A is positive (negative)

b) A is positive (negative) semidefinite, iff all eigenvalues of A are nonnegative (non positive).

c) A is indefinite, iff some of its eigenvalues are positive and some are negative.

Local ~~ex~~ extrema theorem. Suppose that $f(\vec{x})$ is a function with continuous first and second partial derivatives on some set D in $\mathbb{R}^n$. Suppose $\vec{x}^*$ is an interior point of D and that $\vec{x}^*$ is a critical point of $f(\vec{x})$. Then $\vec{x}^*$ is:

a) a strict local minimizer of $f(\vec{x})$ if $Hf(\vec{x}^*)$ is positive definite

b) a strict local maximizer of $f(\vec{x})$ if $Hf(\vec{x}^*)$ is negative definite.

Examples: Minimize

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - x_1 x_2 + x_2 x_3 - x_1 x_3.$$

$$\nabla f(x_1, x_2, x_3) = \left(2x_1 - x_2 - x_3,\ 2x_2 - x_1 + x_3,\ 2x_3 + x_2 - x_1\right)^T$$

Critical points:

$$\begin{aligned} 2x_1 - x_2 - x_3 &= 0 \\ -x_1 + 2x_2 + x_3 &= 0 \\ -x_1 + x_2 + 2x_3 &= 0 \end{aligned}$$

$$\vec{x}^* = (0,0,0)^T$$

$$Hf(\vec{x}) = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}$$

$\Delta_1 > 0$, $\Delta_2 > 0$ & $\Delta_3 > 0$ $\therefore$ Hf is positive definite and $\vec{x}^*$ is a strict global minimizer for $f(\vec{x})$. In fact, since $\vec{x}^*$ is the only critical point and $\nabla f(\vec{x})$ is continuous in $\mathbb{R}^3$, $\vec{x}^*$ is the only strict global minimizer.

b) Minimize

$$f(x,y,z) = e^{x-y} + e^{y-x} + e^{x^2} + z^2$$

$$\nabla f(x,y,z) = \left(e^{x-y} - e^{y-x} + 2x e^{x^2},\ -e^{x-y} + e^{y-x},\ 2z\right)^T$$

If $\nabla f(x,y,z) = \vec{0}$, then $z=0$, $-e^{x-y} + e^{y-x} = 0 \Rightarrow e^{x-y} = e^{y-x} \Rightarrow x-y = y-x \Rightarrow x=y$ and $e^{x-y} - e^{y-x} + 2x e^{x^2} = 0$, but since $x=y$, then $2xe^{x^2} = 0 \Rightarrow x=0$ and $y=0$. So

$$\vec{c}^* = (0,0,0)^T$$

Now,

$$Hf(x,y,z) = \begin{pmatrix} e^{x-y} + e^{y-x} + 4x^2 e^{x^2} + 2x^2 & -e^{x-y} - e^{y-x} & 0 \\ -e^{x-y} - e^{y-x} & e^{x-y} + e^{y-x} & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

$\Delta_1 > 0$ because all the terms of it are positive.

$$\Delta_2 = (e^{x-y}+e^{y-x})(e^{x-y}+e^{y-x}+4x^2e^{x^2}+2x^2) - (e^{x-y}+e^{y-x})^2$$

$$= (e^{x-y}+e^{y-x})^2 + (e^{x-y}+e^{y-x})(4x^2e^{x^2}+2x^2) - (e^{x-y}+e^{y-x})^2$$

$$= (e^{x-y}+e^{y-x})(4x^2e^{x^2}+2x^2) > 0$$

because each factor is positive.

Finally, $\Delta_3 = 2\Delta_2 > 0$.

Therefore, $\vec{c}^*$ is a strict global minimizer for $f(x,y,z)$.

c) Find the global minimizers of

$$f(x,y) = e^{x-y} + e^{y-x}$$

$$\nabla f(x,y) = \begin{pmatrix} e^{x-y} - e^{y-x} \\ -e^{x-y} + e^{y-x} \end{pmatrix}$$

If $\nabla f(x,y) = \vec{0}$, then

$e^{x-y} = e^{y-x} \Rightarrow x-y = y-x \Rightarrow \underbrace{x=y}_{\text{line}}$

So, all points along $y=x$ are critical points.

Now,

$$Hf(x,y) = \begin{pmatrix} e^{x-y}+e^{y-x} & -e^{x-y}-e^{y-x} \\ -e^{x-y}-e^{y-x} & e^{x-y}+e^{y-x} \end{pmatrix}$$

$\Delta_1 > 0$, but $\Delta_2 = 0$. So, by our earlier discussion:

"A is positive semidefinite if of all of its principal minors are non-negative."

All points $x=y$ are global minimizers of $f(x,y)$.