# UNCONSTRAINED OPTIMIZATION : FUNCTIONS OF SEVERAL VARIABLES

1. $f$ IS A FUNCTION OF $n$ VARIABLES $x_1, x_2, \ldots, x_n$

$x$ DENOTES THE VECTOR $(x_1, x_2, \ldots, x_n)$

$f(x) = f(x_1, x_2, \ldots, x_n)$

GRADIENT OF $f$ AT $x$, WRITTEN $\nabla f(x)$, IS THE VECTOR

$$\left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right]^T$$

$\nabla f(x)$ GIVES THE DIRECTION OF STEEPEST ASCENT OF THE FUNCTION $f$ AT POINT $x$.

**EXAMPLE** $f(x_1, x_2) = (x_1 - 2)^2 + 2(x_2 - 1)^2$

$\nabla f(x) = [2(x_1 - 2), 4(x_2 - 1)]^T$ $\square$

2. GRADIENT ACTS LIKE THE DERIVATIVE IN THAT SMALL CHANGES AROUND $x_0$ CAN BE ESTIMATED USING THE GRADIENT.

$$f(x) \simeq f(x_0) + (\nabla f)_{x_0} \cdot (x - x_0)$$

DOT PRODUCT

**DOT PRODUCT** OF VECTORS $(a_1, a_2, a_3)$ AND $(b_1, b_2, b_3)$ IS $(a_1 b_1 + a_2 b_2 + a_3 b_3)$. $a_1, a_2, a_3; b_1, b_2, b_3 \in \mathbb{R}$.

**EXAMPLE** $f(x_1, x_2) = x_1^2 - 3x_1 x_2 + x_2^2$

$f(1, 1) = -1$

WHAT ABOUT $f(1.01, 1.01)$

$$\nabla f(x) = [2x_1 - 3x_2 \ , -3x_1 + 2x_2]^T$$

$$(\nabla f(x))_{(1,1)} = [-1, -1]^T$$

$$(x - x_0) = (0.01, 0.01) \quad ; \quad x_0 = (1,1)$$

$$(\nabla f(x))_{(1,1)} \cdot (x - x_0) = (-0.01 - 0.01) = -0.02$$

$$f(1.01, 1.01) = f(1,1) - 0.02 = -1 - .02 = -1.02$$

DIRECT COMPUTATION

$$f(1.01, 1.01) = (1.01)^2 - 3(1.01)(1.01) + (1.01)^2$$

$$= (1 + .02) - 3(1 + 0.02) + (1 + .02) = -1.02$$

$\square$

## DIRECTIONAL DERIVATIVE $a$ IS A UNIT VECTOR

$\nabla\phi \cdot a$ = DIRECTIONAL DERIVATIVE OF $\phi$ IN THE DIRECTION $a$

= COMPONENT OF $\nabla\phi$ IN THE DIRECTION OF $a$

= RATE OF CHANGE OF $\phi$ AT $(x,y,z)$ IN THE DIRECTION $a$

□

**EXAMPLE** SHOW THAT $\nabla\phi$ IS A VECTOR $\perp^r$ TO THE SURFACE $\phi(x,y,z) = c$, WHERE $c$ IS A CONSTANT.

**SOLUTION** LET $P = (x,y,z)$ BE A POINT ON THE SURFACE. THEN $(dx, dy, dz)$ LIES IN THE TANGENT PLANE TO THE SURFACE AT THE POINT P.

$$0 = d\phi = \frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy + \frac{\partial\phi}{\partial z}dz = \nabla\phi \cdot (dx, dy, dz) = 0$$

DOT PRODUCT

$\Rightarrow$ $\nabla\phi$ IS $\perp^r$ to $(dx, dy, dz)$ AND THEREFORE TO THE SURFACE.

□

## EXAMPLE

CONSIDER THE FUNCTION $f(x_1, x_2) = x_1 \ln x_2$

(a) COMPUTE THE GRADIENT OF $f$

(b) GIVE THE VALUE OF THE FUNCTION $f$ AND GIVE ITS GRADIENT AT THE POINT $(3,1)$

(c) USE THE FORMULA FOR SMALL CHANGES TO OBTAIN AN APPROXIMATE VALUE OF THE FUNCTION AT THE POINT $(2{\cdot}99, 1{\cdot}05)$

## SOLUTION

a) $\nabla f(x) = [\ln x_2 \,,\; x_1/x_2]^T$

b) $f(3,1) = 3 \ln 1 = 0$

$\nabla f(x)\big|_{(3,1)} = [\ln 1, \; 3]^T = [0, 3]^T$

c) $f(2{\cdot}99, 1{\cdot}05) = f(3,1) + (0, 3) \cdot (-{\cdot}01, {\cdot}05)$

$= 0 + {\cdot}15 = 0{\cdot}15$ $\square$

## HESSIAN MATRIX

$$H(x) = \left[ \frac{\partial^2 f}{\partial x_i \partial x_j} \right] \quad \text{IS A } n \times n \text{ MATRIX}$$

$H(x)$ IS A SYMMETRIC MATRIX

**EXAMPLE** FIND HESSIAN MATRIX OF $f(x_1, x_2) = (x_1 - 2)^2 + (x_2 - 1)^2 2$

$$\frac{\partial^2 f}{\partial x_1^2} = 2 \; ; \quad \frac{\partial^2 f}{\partial x_2^2} = 4 \; ; \quad \frac{\partial f}{\partial x_1} = 2(x_1 - 2) \; ; \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0$$

$$H = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

□

## MAXIMUM AND MINIMUM

OPTIMA CAN OCCUR IN THREE PLACES:

1. AT THE BOUNDARY OF THE DOMAIN
2. AT A NONDIFFERENTIAL POINT, OR
3. AT A POINT $x^*$, WHERE $\nabla f(x^*) = 0$

WE IDENTIFY THE FIRST TYPE OF POINT WITH KKT CONDITIONS

THE SECOND TYPE IS FOUND ONLY BY AD HOC METHODS.

THE THIRD TYPE OF POINT CAN BE FOUND BY SOLVING THE

GRADIENT EQUATIONS

## FACTS

LET $\nabla f(x^*) = 0$

1. IF $H(x^*)$ IS POSITIVE DEFINITE, THEN $x^*$ IS A LOCAL MINIMUM
2. IF $H(x^*)$ " NEGATIVE DEFINITE, " " " " " MAXIMUM

## GLOBAL OPTIMA

1. WE SAY THAT A DOMAIN IS CONVEX IF EVERY LINE DRAWN BETWEEN TWO POINTS IN THE DOMAIN LIES WITHIN THE DOMAIN.
2. WE SAY THAT A FUNCTION $f$ IS CONVEX IF THE LINE CONNECTING ANY TWO POINTS LIES ABOVE THE FUNCTION.

   THAT IS, $\forall\, x, y$ IN THE DOMAIN AND $0 < \alpha < 1$, WE HAVE

   $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$.

- IF A FUNCTION IS CONVEX ON A CONVEX DOMAIN, THEN ANY LOCAL MINIMUM IS A GLOBAL MINIMUM
- IF A FUNCTION IS CONCAVE ON A CONVEX DOMAIN, THEN ANY LOCAL MAXIMUM IS GLOBAL MAXIMUM.

- TO CHECK IF A FUNCTION IS CONVEX ON A DOMAIN, CHECK THAT ITS HESSIAN MATRIX $H(x)$ IS POSITIVE SEMIDEFINITE FOR EVERY POINT $x$ IN THE DOMAIN

  TO CHECK IF A FUNCTION IS CONCAVE ON A DOMAIN, CHECK THAT ITS HESSIAN MATRIX $H(x)$ IS NEGATIVE SEMIDEFINITE FOR EVERY POINT $x$ IN THE DOMAIN

EXAMPLE FIND THE LOCAL EXTREMA OF $f(x_1, x_2) = x_1^3 + x_2^3 - 3x_1x_2$

SOLUTION THE FUNCTION IS EVERYWHERE DIFFERENTIABLE

SO EXTREMA CAN OCCUR ONLY AT POINTS $x^*$ SUCH THAT $\nabla f(x^*) = 0$

$$\nabla f(x) = \left[3x_1^2 - 3x_2 \quad 3x_2^2 - 3x_1\right]^T$$

$$= 3\left[x_1^2 - x_2 \quad x_2^2 - x_1\right]^T$$

$\nabla f(x)$ IS ZERO AT $(x_1, x_2) = (0, 0)$ OR $(x_1, x_2) = (1, 1)$

HESSIAN $H(x)$ IS

$$H(x) = \begin{bmatrix} 6x_1 & -3 \\ -3 & 6x_2 \end{bmatrix}$$

a) $H(0,0) = \begin{bmatrix} 0 & -3 \\ -3 & 0 \end{bmatrix}$

LET $H_1$ BE FIRST PRINCIPAL MINOR OF $H(0,0) = [0]$; $\det H_1 = 0$

" $H_2$ " SECOND " " " $H(0,0) = H(0,0)$; $\det H_2 = -9 < 0$

$\therefore$ $H(0,0)$ IS NEITHER POSITIVE NOR NEGATIVE DEFINITE

b) $H(1,1) = \begin{bmatrix} 6 & -3 \\ -3 & 6 \end{bmatrix}$

LET $H_1$ BE THE FIRST PRINCIPAL MINOR OF $H(1,1) = [6]$; $\det H_1 = 6 > 0$

" $H_2$ " " SECOND " " " $H(1,1) = H(1,1)$; $\det H_2 = 27 > 0$

$\therefore$ $H(1,1)$ IS POSITIVE DEFINITE $\Rightarrow$ $(1,1)$ IS A LOCAL MINIMUM. $\square$

**EXAMPLE** FIND THE LOCAL EXTREMA OF $f(x_1, x_2, x_3) = x_1^2 + (x_1+x_2)^2 + (x_1+x_3)^2$

**SOLUTION**

$$\nabla f(x) = [2x_1 + 2(x_1+x_2) + 2(x_1+x_3),\ 2(x_1+x_2),\ 2(x_1+x_3)]^T$$
$$= [0, 0, 0]^T \Rightarrow x_1 = x_2 = x_3 = 0.$$

$$H(x) = \begin{bmatrix} 6 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$

DETERMINANTS OF PRINCIPAL MINOR ARE:

$\det H_1 = 6$; $\det H_2 = 8$; $\det H_3 = 24 - 2\cdot 4 + 2(-4) = 8$

ALL THESE DETERMINANTS ARE POSITIVE.

SO $H(0,0,0)$ IS POSITIVE DEFINITE, AND $x_1 = x_2 = x_3 = 0$ IS A MINIMUM. $\square$

**EXAMPLE** SHOW THAT $f(x_1, x_2, x_3) = x_1^4 + (x_1+x_2)^2 + (x_1+x_3)^2$ IS CONVEX OVER $\mathbb{R}^3$.

**SOLUTION**

$$\frac{\partial f}{\partial x_1} = 4x_1^3 + 2(x_1+x_2) + 2(x_1+x_3)$$

$$\frac{\partial f}{\partial x_2} = 2(x_1+x_2) \ ; \quad \frac{\partial f}{\partial x_3} = 2(x_1+x_3)$$

$$H(x_1, x_2, x_3) = \begin{bmatrix} 12x_1^2+4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$

THE DETERMINANTS OF PRINCIPAL MINORS ARE:

$\det H_1 = 12x_1^2 + 4 > 0$; $\det H_2 = 2(12x_1^2+4) - 4 = 4(6x_1^2+1) > 0$

$\det H_3 = (12x_1^2+4)4 - 2\cdot 4 - 8 = 48x_1^2 \geq 0$

SO $H(x_1, x_2, x_3)$ IS POSITIVE SEMIDEFINITE FOR $\forall\ (x_1, x_2, x_3) \in \mathbb{R}^3$.

$\Rightarrow f$ IS CONVEX OVER $\mathbb{R}^3$. $\square$

# CONVEXITY / CONCAVITY OF TWO-VARIABLE FUNCTION

LET $f(\cdot,\cdot)$ BE A FUNCTION OF VARIABLES $x_1, x_2 \in \mathbb{R}$

ASSUME THAT THE DERIVATIVES EXIST.

LET $\dfrac{\partial^2 f(x_1,x_2)}{\partial x_1^2} \triangleq f_{x_1,x_1}$ ; $\dfrac{\partial^2 f(x_1,x_2)}{\partial x_1 \partial x_2} \triangleq f_{x_1,x_2}$ ; $\dfrac{\partial^2 f(x_1,x_2)}{\partial x_2^2} \triangleq f_{x_2,x_2}$

HESSIAN $= H(x_1,x_2) = \begin{bmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_1x_2} & f_{x_2x_2} \end{bmatrix}$ ; DETERMINANT OF $H(x_1,x_2) \triangleq |H(x_1,x_2)|$

$f(x_1,x_2)$ IS CONVEX IF:

i) $|H(x_1,x_2)| \geq 0$

ii) $f_{x_1,x_1} \geq 0$

OR

iii) $f_{x_2,x_2} \geq 0$

det of H

CONCAVE IF

i) $|H(x_1,x_2)| \geq 0$

ii) $f_{x_1,x_1} \leq 0$

OR

iii) $f_{x_2x_2} \leq 0$

# STATEMENTS ABOUT MAXIMA & MINIMA OF TWO-VARIABLE FUNCTION

LET $f(\cdot,\cdot)$ BE A FUNCTION OF TWO VARIABLES $x_1, x_2 \in \mathbb{R}$

ASSUME THAT THE DERIVATIVES EXIST

LET $\nabla f(x_1, x_2) = 0$ AT $(x_1^*, x_2^*)$

ALSO LET $\det H(x_1, x_2) = f_{x_1 x_1} f_{x_2 x_2} - \left(f_{x_1 x_2}\right)^2$

$$\det H(x_1^*, x_2^*) = E$$

$$f_{x_1 x_1}(x_1^*, x_2^*) = f_{11} \; ; \; f_{x_2 x_2}(x_1^*, x_2^*) = f_{22}$$

$$f_{x_1 x_2}(x_1^*, x_2^*) = f_{12}$$

## STATEMENT ABOUT LOCAL EXTREMUM

1. IF $E > 0$, AND $f_{11} > 0$ (HENCE $f_{22} > 0$); WE HAVE A MINIMUM
2. IF $E > 0$, AND $f_{11} < 0$ (HENCE $f_{22} < 0$); WE HAVE A MAXIMUM
3. IF $E < 0$, AND $f_{11} f_{22} \neq 0$, THERE IS NO MAXIMUM OR MINIMUM BUT A SADDLE POINT.
4. IF $E = 0$, OR IF $f_{11} f_{22} = 0$, WE HAVE TO EXAMINE THE HIGHER DERIVATIVES OR INVESTIGATE THE FUNCTIONAL VALUES AT AND NEAR $(x_1^*, x_2^*)$