NUMERICAL METHODS

NUMERICAL METHODS FOR MINIMIZING A FUNCTION.

- NEWTON'S METHOD
- GRADIENT DESCENT

MAXIMIZING PROBLEMS CAN BE CONVERTED TO MINIMIZATION EASILY. FOR EXAMPLE: MAX $f(x) \Leftrightarrow$ MIN $-f(x)$

NEWTON'S METHOD FOR MINIMIZING A FUNCTION OF A SINGLE VARIABLE.

NEWTON'S METHOD FOR MINIMIZING A FUNCTION OF A SINGLE VARIABLE USES A QUADRATIC APPROXIMATION OF FUNCTION $f(x)$, WHERE $x \in \mathbb{R}$

i) IT ASSUMES THAT $f(\cdot)$ IS TWICE DIFFERENTIABLE.

ii) ALSO ASSUME THAT $f(\cdot)$ IS CONVEX. (STRICTLY)

DENOTE THE FIRST AND SECOND DERIVATIVE OF $f(x)$ BY $f'(x)$ AND $f''(x)$ RESPECTIVELY. (WITH RESPECT TO x)

USE TAYLOR'S EXPANSION OF $f(x)$ AROUND x_0.

$$f(x) \cong f(x_0) + (x-x_0)f'(x_0) + \frac{(x-x_0)^2}{2} f''(x_0).$$

MINIMIZE $f(\cdot)$ BY TAKING DERIVATIVE ON BOTH SIDES WITH RESPECT TO x. THIS LEADS TO:

$$f'(x) = f'(x_0) + (x-x_0)f''(x_0)$$

LET $f'(x) = 0 \Rightarrow \boxed{x = x_0 - \frac{f'(x_0)}{f''(x_0)}}$

NOTE THAT $f''(x_0) > 0$ (BECAUSE OF STRICT CONVEXITY ASSUMPTION)

THE ABOVE EQUATION CAN BE USED TO UPDATE x UNTIL IT CONVERGES TO THE LOCATION OF THE OPTIMUM VALUE.

ALGORITHM

$\epsilon > 0$

$x_0 \leftarrow$ INITIAL VALUE

WHILE $|f'(x_0)| > \epsilon$ DO

$$x_0 \leftarrow x_0 - \frac{f'(x_0)}{f''(x_0)}$$

END WHILE

RETURN x_0

MINIMIZATION OF A FUNCTION OF SEVERAL VARIABLES

$f : \mathbb{R}^n \rightarrow \mathbb{R}$

$x \in \mathbb{R}^n$; $f(x) \in \mathbb{R}$

i) ASSUME THAT $f(\cdot)$ IS TWICE DIFFERENTIABLE

ii) ASSUME THAT $f(\cdot)$ IS STRICTLY CONVEX AND HENCE ITS HESSIAN IS POSITIVE DEFINITE AND INVERTIBLE.

USE GENERALIZED TAYLOR'S EXPANSION OF $f(\cdot)$ AT $x_0 \in \mathbb{R}^n$.

$$f(x) \simeq f(x_0) + (x - x_0)^T \nabla f(x_0) + \frac{1}{2}(x - x_0)^T H(x_0)(x - x_0)$$

WHERE $\nabla f(x)$ IS THE GRADIENT OF $f(x)$, &

$H(x)$ IS THE HESSIAN MATRIX

MINIMIZE $f(x)$ BY TAKING GRADIENT ON BOTH SIDES WITH RESPECT TO x, AND SETTING IT TO ZERO. THAT IS

$$\nabla f(x) = \nabla f(x_0) + H(x_0)(x - x_0)$$

$$\nabla f(x) = 0 \Rightarrow \boxed{x = x_0 - H^{-1}(x_0)\nabla f(x_0)}$$

NOTE THAT $H^{-1}(x_0)$ EXISTS BECAUSE OF THE ASSUMPTION OF STRICT CONVEXITY

THE ABOVE EQUATION CAN BE USED TO UPDATE x UNTIL IT CONVERGES TO THE LOCATION OF THE OPTIMUM VALUE.

METHOD OF STEEPEST DESCENT

- SOME TIMES IT IS NOT POSSIBLE TO FIND THE MINIMUM OF A FUNCTION ANALYTICALLY
- OCCASSIONALLY NEWTON'S NUMERICAL TECHNIQUE MIGHT BE UNRELIABLE.
- USE METHOD OF STEEPEST DESCENT TO FIND MINIMUM OF A FUNCTION NUMERICALLY.
- GIVEN $f : \mathbb{R}^n \longrightarrow \mathbb{R}$; WHERE $\mathbb{R} = (-\infty, +\infty)$

f IS DIFFERENTIABLE AT POINT x_0

THE DIRECTION OF STEEPEST DESCENT IS THE DIRECTION OF THE VECTOR $-\nabla f(x)\Big|_{x=x_0}$ (AT POINT $x_0 \in \mathbb{R}^n$)

ASIDE: DOT PRODUCT OF TWO VECTORS

LET $a, b \in \mathbb{R}^3$; $\|\cdot\|$ = EUCLIDEAN NORM (USUAL DISTANCE MEASURE)

LET $a = (a_1, a_2, a_3)$, & $b = (b_1, b_2, b_3)$

$$\|a\|^2 = a_1^2 + a_2^2 + a_3^2 \ ; \ \|b\|^2 = b_1^2 + b_2^2 + b_3^2$$

$$a \cdot b = a_1 b_1 + a_2 b_2 + a_3 b_3$$

↖ DOT PRODUCT OPERATOR

$$= \|a\| \|b\| \cos\theta$$

θ = ANGLE BETWEEN VECTORS a & b

□

NOTE: $|a \cdot b| \leq \|a\| \|b\|$

DISCUSSION IT IS ESTABLISHED THAT THE DIRECTION OF STEEPEST DESCENT ; IS THE DIRECTION OF THE VECTOR $-\nabla f(x_0)$, AT POINT $x_0 \in \mathbb{R}^n$

LET $u \in \mathbb{R}^n$ BE A UNIT VECTOR, THEN $\|u\| = 1$.

THE RATE OF CHANGE OF f AT x IN THE DIRECTION OF u IS

$\nabla f(x) \cdot u$

NOTE THAT: $|\nabla f(x) \cdot u| \leq \|\nabla f(x)\| \cdot \|u\|$

THE UPPER-BOUND IS ACHIEVABLE, IF u IS PARALLEL TO $\nabla f(x)$. THAT IS, WHEN $u = \nabla f(x) / \|\nabla f(x)\|$

HOWEVER, WE WANT TO MOVE IN A DIRECTION IN WHICH $f(x)$ IS MINIMIZED. THEREFORE, WE SET

$$u = -\frac{\nabla f(x)}{\|\nabla f(x)\|}$$

$\square$

ITERATIVE STEPS OF THE METHOD OF STEEPEST DESCENT

GIVEN A POINT $x^{(k)} \in \mathbb{R}^n$, THE NEXT POINT $x^{(k+1)} \in \mathbb{R}^n$ IS COMPUTED AS FOLLOWS:

1. COMPUTE $\nabla f(x^{(k)})$

2. SET $\phi_k(t) = f(x^{(k)} - t\,\nabla f(x^{(k)}))$.

THAT IS, ϕ_k EVALUATES f ALONG THE LINE THROUGH $x^{(k)}$ IN THE DIRECTION OF STEEPEST DESCENT.

3. LET t_k BE THE GLOBAL MINIMIZER OF $\phi_k(t)$.

THIS t_k TELLS US HOW FAR ALONG THE LINE WE WANT TO GO.

4. GO THAT FAR ALONG THE LINE: SET

$$x^{(k+1)} = x^{(k)} - t_k \nabla f(x^{(k)})$$

$\square$

EXAMPLE APPLY METHOD OF STEEPEST DESCENT TO THE FUNCTION

$$f(x,y) = 4x^2 - 4xy + 2y^2$$

WITH INITIAL POINT $x_0 = (2,3)$

SOLUTION

$$\nabla f(x,y) = (8x-4y, -4x+4y)$$

$$\nabla f(x_0) = \nabla f(2,3) = (4,4)$$

STEP 1: MINIMIZE THE FUNCTION

$$\varphi_0(t) = f(x_0 - t\nabla f(x_0))$$

$$= f((2,3) - t(4,4)) = f(2-4t, 3-4t)$$

ALSO $\varphi_0'(t) = \nabla f(2-4t, 3-4t) \cdot (-4,-4)$

$$= (8(2-4t) - 4(3-4t), -4(2-4t) + 4(3-4t)) \cdot (-4,-4)$$

$$= (4-16t, 4) \cdot (-4,-4) = -16(1-4t+1)$$

$$= (64t - 32)$$

$\varphi_0'(t) = 0 \Rightarrow t = 1/2$; $\varphi_0''(t) = 64 > 0$

THAT IS $\varphi_0(t)$ HAS A GLOBAL MINIMA AT $t = 1/2$. THUS

$$x_1 = x_0 - t_0 \nabla f(x_0) = (2,3) - \frac{1}{2}(4,4) = (0,1)$$

STEP 2: MINIMIZE THE FUNCTION

$$\varphi_1(t) = f(x_1 - t\nabla f(x_1))$$

$$= f((0,1) - t(-4,+4)) = f(4t, 1-4t)$$

BY COMPUTING

$$\varphi_1'(t) = \nabla f(4t, 1-4t) \cdot (4,-4)$$
$$= (8 \cdot 4t - 4(1-4t), -4(4t) + 4(1-4t)) \cdot (4,-4)$$
$$= (48t - 4, -32t + 4) \cdot (4,-4)$$
$$= 16(12t - 1 + 8t - 1)$$
$$= 320t - 32$$

$\varphi_1'(t) = 0 \Rightarrow t = 1/10$; $\varphi''(t) = 320 > 0$

THAT IS $\varphi_1(t)$ HAS A GLOBAL MINIMA AT $t = 1/10$. THUS

$$x_2 = x_1 - t_1 \nabla f(x_1) = (0,1) - \tfrac{1}{10}(-4,4)$$
$$= (0,1) + \left(\tfrac{2}{5}, -\tfrac{2}{5}\right) = \left(\tfrac{2}{5}, \tfrac{3}{5}\right)$$

STEP 3: $x_3 = \left(0, \frac{3}{10}\right)$. WHY?

COMPLETE THIS STEP AS AN ASSIGNMENT PROBLEM. OBSERVE THAT THE METHOD OF STEEPEST DESCENT PRODUCES A SEQUENCE OF ITERATES x_k THAT IS CONVERGING TO THE STRICT GLOBAL MINIMIZER OF $f(x,y)$ AT $x^* = (0,0)$.

$\square$

SOLUTION VIA ANALYTICAL MEANS.

MINIMIZE $f(x,y) = 4x^2 - 4xy + 2y^2$

$\frac{\partial f}{\partial x} = 8x - 4y$; $\frac{\partial^2 f}{\partial x^2} = 8$

$\frac{\partial f}{\partial y} = -4x + 4y$; $\frac{\partial^2 f}{\partial y^2} = 4$; $\frac{\partial^2 f}{\partial x \partial y} = -4$

EXTREMA CAN OCCUR AT $\nabla f(x,y) = 0$

WE HAVE $8x - 4y = 0$

$-4x + 4y = 0 \Rightarrow x = y$

$0 = 8x - 4y = 8x - 4x = 4x \Rightarrow x = 0$

THEREFORE EXTREMA CAN OCCUR AT $(x,y) = (0,0)$

HESSIAN MATRIX $= H(x,y)$

$$= \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 8 & -4 \\ -4 & 4 \end{bmatrix}$$

i) FIRST PRINCIPAL MINOR OF $H(0,0) = [8]$

ii) SECOND " " " $= \det H(0,0) = 32 - 16 = 16 > 0$

i) & ii) $\Rightarrow$ $(0,0)$ IS A LOCAL MINIMA

$f(x,y)$ IS ALSO CONVEX.

$\square$