Taylor's Formula

This Hypercard stack was prepared by:
Dennis L. Bricker,
Dept. of Industrial Engineering,
University of Iowa,
Iowa City, Iowa 52242
e-mail: dbricker@icaen.uiowa.edu

author

Function of One Variable

Suppose that f(x), f'(x), and f''(x) exist on the closed interval $[a,b] = \{x \;\; R \mid a \le x \le b\}$.
If x* and x are any two distinct points in [a,b], then there exists a point z between x* and x such that

$$f(x) = f(x^*) + f'(x^*)(x - x^*) + \frac{f''(z)}{2}(x - x^*)^2$$

Taylor's Formula

$$f(x) = f(x^*) + f'(x^*)(x - x^*) + \frac{f''(z)}{2}(x - x^*)^2$$

If $f''(x) > 0$ for all x, and $f'(x^*) = 0$, then Taylor's Formula tells us that

$$f(x) = f(x^*) + 0 + \text{a positive number} > f(x^*)$$

That is, x^* is the point that minimizes the function f.

Critical Point

The point x^* is a *critical point* of a function f if $f'(x^*)$ exists and equals zero.

(stationary point)

Function of Several Variables

Gradient

vector of first partial derivatives

$$\nabla \mathbf{f}(\mathbf{x}) = \left[\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \cdots \frac{\partial f(x)}{\partial x_n}\right]$$

Hessian

matrix of second partial derivatives

$$\nabla^2 \mathbf{f}(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

Function of Several Variables

Suppose that x* and x are points in R and that f(x) is a function of n variables with continuous first and second partial derivatives on some open set containing the line segment [x*,x] joining x* and x. Then there exists a $z \in [x^*,x]$ such that

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}^*) + \nabla \mathbf{f}(\mathbf{x}^*) \cdot (\mathbf{x} - \mathbf{x}^*) + \frac{1}{2} (\mathbf{x} - \mathbf{x}^*) \cdot \nabla^2 \mathbf{f}(\mathbf{z}) \, (\mathbf{x} - \mathbf{x}^*)$$

Taylor's Formula

QUADRATIC FORM

$$f(x_1, x_2, \cdots x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_i^j x_i x_j = x^T A x$$

A is not unique, but we can choose A to be symmetric $\left(A = \frac{1}{2}\nabla^2 f(x)\right)$

$$x_1^2 + x_1x_2 + 3x_2^2 = [x_1 \ \ x_2]\begin{bmatrix} 1 & 1/2 \\ 1/2 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= [x_1 \ \ x_2]\begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

A_i^i = coefficient of x_i^2

A_i^j = 1/2 of coefficient of $x_i x_j$

Which are quadratic forms?

$x_1 + 2x_2^2$

x_1x_2

$3x_1^2 - x_1x_2$

$x_1x_2 - x_2x_3 + x_1x_3$

$$x_1^2 + x_2^2 = x^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x$$

>0 for $x \neq 0$

positive definite

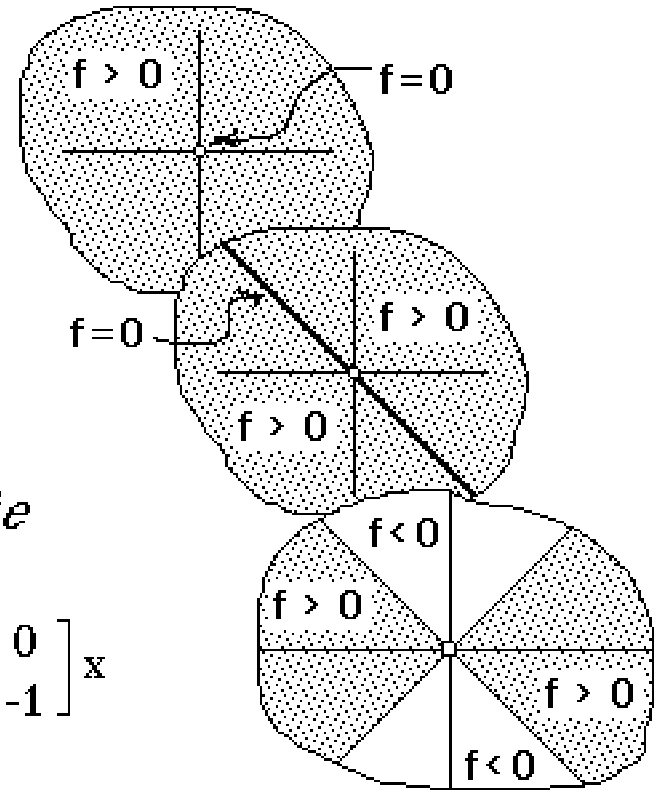
$$x_1^2 + 2x_1x_2 + x_2^2 = x^T \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} x$$

$= (x_1 + x_2)^2 \geq 0$ for all x

positive semidefinite

$$x_1^2 - x_2^2 = x^T \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} x$$

indefinite



Positive Definite a square symmetric matrix A is positive definite if

$$x^t A x > 0 \text{ for all } x \neq 0$$

Note: a symmetric matrix whose entries are all positive need not be positive definite.

Example: $A = \begin{bmatrix} 1 & 4 \\ 4 & 1 \end{bmatrix}$

Let $x=[1,-1]$: $\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -6 < 0$

Positive Definite

A symmetric matrix with some negative elements may be positive definite.

Example: $A = \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix}$

$$x^t A x = x_1^2 - 2x_1x_2 + 4x_2^2 = (x_1 - x_2)^2 + 3x_2^2 > 0$$

for all $x \neq 0$

Positive Semidefinite

a square symmetric matrix A is positive semidefinite if $x^t Ax \geq 0$ for all x

.
.
.

Negative Definite

A square symmetric matrix A is negative definite if

$$x^t A x < 0 \quad \text{for all} \quad x \neq 0$$

Negative Semidefinite

A square symmetric matrix A is negative semidefinite if $x^t A x \leq 0$ for all x

Indefinite

A square symmetric matrix A is indefinite if

$$\exists\, x^+ \text{ such that } (x^+)^t A\, x^+ > 0,$$

and

$$\exists\, x^- \text{ such that } (x^-)^t A\, x^- < 0$$

i.e., if it is neither positive semidefinite nor negative semidefinite.

Diagonal Matrices

A diagonal matrix D is

- positive definite if $D_i^i > 0$ for all i
- positive semidefinite if $D_i^i \geq 0$ for all i
- negative definite if $D_i^i < 0$ for all i
- negative semidefinite if $D_i^i \leq 0$ for all i

$$x^t D\, x = \sum_{i=1}^{n} D_i^i\, x_i^2$$

Suppose that a symmetric matrix A is reduced to upper triangular form by use of the elementary row operation

- Add to any row a scalar multiple of another row

without using

- Multiply any row of the matrix by a (positive or negative) scalar
- Interchange two rows of the matrix

Then A is

- positive definite if $U_i^i > 0 \;\; \forall\, i$
- positive semidefinite if $U_i^i \geq 0 \;\; \forall\, i$
- negative definite if $U_i^i < 0 \;\; \forall\, i$
- negative semidefinite if $U_i^i \leq 0 \;\; \forall\, i$

WHY?

Consider the quadratic form $\mathbf{x}^T \mathbf{A}\,\mathbf{x} = \sum_i^n \sum_j^n A_i^j x_i x_j$

$$\mathbf{x}^T \mathbf{A}\,\mathbf{x} = \mathbf{x}^T \mathbf{L}\mathbf{D}\mathbf{L}^T \mathbf{x} = \left[\mathbf{L}^T \mathbf{x}\right]^T \mathbf{D} \left[\mathbf{L}^T \mathbf{x}\right] = \mathbf{y}^T \mathbf{D}\,\mathbf{y} = \sum_i^n D_i^i y_i^2$$

where $\mathbf{y} = \mathbf{L}^T \mathbf{x}$

If $D_i^i \geq 0$, then, $\mathbf{x}^T \mathbf{A}\,\mathbf{x} \geq 0$ for all x

A is positive semidefinite

If $D_i^i > 0$, $\mathbf{x}^T \mathbf{A}\,\mathbf{x} > 0$ for all $x \neq 0$ ($\Longrightarrow y \neq 0$)

A is positive definite

etc.