Taylor's Series & Quadratic Forms

Useful in forming linear & quadratic approximations of functions

Dept of Mechanical & Industrial Engineering
The University of Iowa

Suppose that the function $f: R^1 \to R^1$ has first & second derivatives. Then

$$f(x) = f(x^0) + f'(x^0)(x-x^0) + \frac{1}{2} f''(x^0)(x-x^0)^2 + \ldots$$

$$= f(x^0) + f'(x^0)(x-x^0) + \frac{1}{2} f''(z)(x-x^0)^2$$

for *some* $z \in (x^0, x)$

Equivalently, letting $x = x^0 + d$:

$$f(x^0+d) = f(x^0) + f'(x^0)d + \frac{1}{2} f''(z) d^2$$

where $z = x^0 + \alpha d$, for some $0 < \alpha < 1$.

functions of single variable

If $f''(x) > 0 \ \forall x$ & $f'(x^0) = 0$,

i.e., the *first* derivative is zero at x^0 and
the *second* derivative is positive everywhere,

then

$$f(x) = f(x^0) + 0 + \frac{1}{2} f''(z)(x-x^0)^2 > f(x^0)$$

That is, x^0 is a strict *minimizer* of the function f.

Definition: The point x* is a ***critical point*** of a function f if f is differentiable at x* and $f'(x^*) = 0$.

Taylor's Formula for functions of multiple variables

$$f(x) = f(x^0) + (x-x^0)\nabla f(x^0) + \frac{1}{2}(x-x^0)\nabla^2 f(z)(x-x^0)$$

$$f(x^0+d) = f(x^0) + d^T \nabla f(x^0) + \frac{1}{2} d^T \nabla^2 f(z) d$$

for some $z = \lambda x^0 + (1-\lambda)x$ where $\lambda \in (0,1)$.

Gradient
vector of first partial derivatives

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

Hessian
matrix of second partial derivatives

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

Example: Quadratic Approximation of a function

Consider $f\left(x_1,x_2\right)=e^{2x_1+3x_2}$.

$$\nabla f\left(x_1,x_2\right)=\begin{bmatrix}2e^{2x_1+3x_2}\\ 3e^{2x_1+3x_2}\end{bmatrix}$$

and

$$\nabla^2 f\left(x_1,x_2\right)=\begin{bmatrix}4e^{2x_1+3x_2} & 6e^{2x_1+3x_2}\\ 6e^{2x_1+3x_2} & 9e^{2x_1+3x_2}\end{bmatrix}$$

Let $x^0=(2,1)$. Then

$$f\left(x^0\right)=e^7,\ \nabla f\left(x^0\right)=\begin{bmatrix}2e^7\\ 3e^7\end{bmatrix},\ \&\ \nabla^2 f\left(x^0\right)=\begin{bmatrix}4e^7 & 6e^7\\ 6e^7 & 9e^7\end{bmatrix}$$

Approximation by Taylor Series:

$$f\left(x_1,x_2\right)\approx e^7+\left[2e^7,3e^7\right]\begin{bmatrix}x_1-2\\ x_2-1\end{bmatrix}+\frac{1}{2}\left[x_1-2,x_2-1\right]\begin{bmatrix}4e^7 & 6e^7\\ 6e^7 & 9e^7\end{bmatrix}\begin{bmatrix}x_1-2\\ x_2-1\end{bmatrix}$$

At $x^0=(2,1)$, the approximation is ***exact***, i.e.,

$$f(2,1)\approx e^7+\left[2e^7,3e^7\right]\begin{bmatrix}0\\ 0\end{bmatrix}+\frac{1}{2}[0,0]\begin{bmatrix}4e^7 & 6e^7\\ 6e^7 & 9e^7\end{bmatrix}\begin{bmatrix}0\\ 0\end{bmatrix}=e^7$$

If $d^T\nabla^2 f(x)d>0\quad\forall x\ \&\ d\neq 0\quad \&\quad \nabla f\left(x^0\right)=0$,

i.e., the *Hessian* matrix $\nabla^2 f(x)$ is *positive definite* everywhere

and

the *gradient* $\nabla f(x)$ is zero at x^0,

then

$$f(x)=f\left(x^0\right)+0+\frac{1}{2}\left(x-x^0\right)\nabla^2 f(z)\left(x-x^0\right)>f\left(x^0\right)$$

That is, $f\left(x^0\right)<f(x)$ if $x\neq x^0$

so that x^0 is a strict *minimizer* of f.

Quadratic Form

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

For a given quadratic form, the matrix A is not uniquely determined, but we can choose A to be the unique symmetric matrix $A=\frac{1}{2}\nabla^2 f(x)$.

EXAMPLE:

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

$$=\left[x_1,x_2\right]^T\begin{bmatrix}1 & 1\\ 0 & 3\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

Notation:
A_i^j = element of matrix A in row i & column j

Which of the following are quadratic forms?

$$x_1 + 2x_2^2 \overset{?}{=} x^T \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} x$$

$$3x_1^2 - x_1 x_2 \overset{?}{=} x^T \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} x$$

$$x_1 x_2 \overset{?}{=} x^T \begin{bmatrix} \_\_ & \_\_ \\ \_\_ & \_\_ \end{bmatrix} x$$

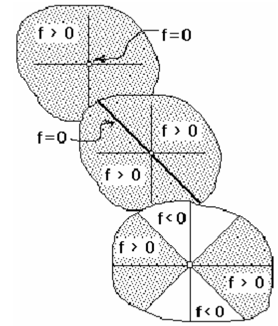
$$x_1 x_2 - x_2 x_3 + x_1 x_3 \overset{?}{=} x^T \begin{bmatrix} \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \\ \_\_ & \_\_ & \_\_ \end{bmatrix} x$$

A square symmetric matrix A is

- ***positive definite*** if $x^T A x > 0 \;\; \forall x \neq 0$
- ***positive semidefinite*** if $x^T A x \geq 0 \;\; \forall x$

Examples

- $x_1^2 + x_2^2 = x^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x$ is ***positive definite*** *(pd)*
- $(x_1 + x_2)^2 = x_1^2 + 2x_1 x_2 + x_2^2 = x^T \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} x$ is ***positive semidefinite*** *(psd)*
- $(x_1 - x_2)(x_1 + x_2) = x_1^2 - x_2^2 = x^T \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} x$ is ***indefinite***



Example: a symmetric matrix whose entries are all positive need not be positive definite!

Consider the matrix $A = \begin{bmatrix} 1 & 4 \\ 4 & 1 \end{bmatrix}$

Select $x = [1, -1]$. Then $[1, -1]\begin{bmatrix} 1 & 4 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = -6 < 0$ ***negative!***

Example:

A matrix with some negative elements *may* be positive definite!

Consider the matrix

$$A=\begin{bmatrix}1 & -1\\ -1 & 4\end{bmatrix}\Rightarrow x^T Ax = x_1^2 - 2x_1x_2 + 4x_2^2 = \left(x_1 - x_2\right)^2 + 3x_2^2 > 0 \quad \forall x \neq 0$$

A square symmetric matrix A is

- ***negative definite*** if $x^T Ax < 0 \quad \forall x \neq 0$
- **negative semidefinite** if $x^T Ax \leq 0 \quad \forall x$

A square symmetric matrix A is ***indefinite*** if

$\exists x^+ \neq 0$ such that $\left(x^+\right)^T Ax^+ > 0$

and

$\exists x^- \neq 0$ such that $\left(x^-\right)^T Ax^- < 0$

i.e., A is neither positive semidefinite nor negative semidefinite!

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

$$D=\begin{bmatrix}D_1^1 & 0 & 0 & \cdots 0\\ 0 & D_2^2 & 0 & \cdots 0\\ 0 & 0 & D_3^3 & \cdots 0\\ 0 & 0 & 0 & \cdots D_n^n\end{bmatrix}$$

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

Testing for Positive Definiteness

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

$$U = \begin{bmatrix} U_1^1 & U_1^2 & U_1^3 & \cdots U_1^n \\ 0 & U_2^2 & U_2^3 & \cdots U_2^n \\ 0 & 0 & U_3^3 & \cdots U_3^n \\ 0 & 0 & 0 & \cdots U_n^n \end{bmatrix}$$

Then A is

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

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.