OPTIMUM DESIGN

Dr. / Ahmed Nagib Elmekawy

Lecture 6

Fall 2017

Taylor Series Expansion

Assume *f*(x) is:

1. Continuous function of a single variable x
2. Differentiable n times
3. x ∈ *S*, where *S* is non-empty, closed, and bounded
4. therefore x* is a possible optima

$$f(x) = f(x^*) + \frac{df(x^*)}{dx}(x - x^*) + \frac{1}{2}\frac{d^2 f(x^*)}{dx^2}(x - x^*)^2 + \frac{1}{3!}\frac{d^3 f(x^*)}{dx^3}(x - x^*)^3 + h.o.t.$$

Taylor Series Approximations

let $x - x^* = d$, a small change or vicinity of x^*...then

$$f(x) = f(x^* + d)$$

$$f(x) = f(x^*) + \frac{df(x^*)}{dx} d + \frac{1}{2} \frac{d^2 f(x^*)}{dx^2} d^2 + \frac{1}{3!} \frac{d^3 f(x^*)}{dx^3} d^3 + h.o.t.$$

$$f(x) \approx f(x^*) + \frac{df(x^*)}{dx} d \qquad \text{First Order Approximation}$$

$$f(x) \approx f(x^*) + \frac{df(x^*)}{dx} d + \frac{1}{2} \frac{d^2 f(x^*)}{dx^2} d^2 \qquad \text{Second Order Approximation}$$

Taylor Series – Mult. Variables

Estimate $f(\mathbf{x})$ using Taylor Series

$$f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f^T(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^T \mathbf{H}(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + R$$

Let $\mathbf{d} = (\mathbf{x} - \mathbf{x}^*)$, therefore $\mathbf{x} = \mathbf{x}^* + \mathbf{d}$

$$f(\mathbf{x}) = f(\mathbf{x}^* + \mathbf{d}) = f(\mathbf{x}^*) + \nabla f^T(\mathbf{x}^*)\mathbf{d} + \frac{1}{2}\mathbf{d}^T \mathbf{H}\,\mathbf{d} + R$$

Change in function

Let's define a change in the function with the symbol Δ

$$\Delta f = f(\mathbf{x}) - f(\mathbf{x}^*)$$

$$\Delta f = \nabla f^T \mathbf{d} + \frac{1}{2}\mathbf{d}^T \mathbf{H}\,\mathbf{d} + R$$

Δ-Delta ∇-Del

Let's define a first order change in the function

$$\delta f = \nabla f^T \delta\mathbf{x} = \nabla f \bullet \delta\mathbf{x}$$

EXAMPLE 4.8 TAYLOR'S EXPANSION OF A FUNCTION OF TWO VARIABLES

Obtain a second-order Taylor's expansion for the function $f(\mathbf{x}) = 3x_1^3 x_2$ at the point $\mathbf{x}^* = (1, 1)$.

Solution

The gradient and Hessian of the function $f(\mathbf{x})$ at the point $\mathbf{x}^* = (1,1)$ using Eqs. (4.5) and (4.8) are

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 9x_1^2 x_2 \\ 3x_1^3 \end{bmatrix} = \begin{bmatrix} 9 \\ 3 \end{bmatrix}; \quad \mathbf{H} = \begin{bmatrix} 18x_1 x_2 & 9x_1^2 \\ 9x_1^2 & 0 \end{bmatrix} = \begin{bmatrix} 18 & 9 \\ 9 & 0 \end{bmatrix} \tag{a}$$

Substituting these in the matrix form of Taylor's expression given in Eq. (4.12), and using $\mathbf{d} = \mathbf{x} - \mathbf{x}^*$, we obtain an approximation $\bar{f}(\mathbf{x})$ for $f(\mathbf{x})$ as

$$\bar{f}(\mathbf{x}) = 3 + \begin{bmatrix} 9 \\ 3 \end{bmatrix}^T \begin{bmatrix} (x_1 - 1) \\ (x_2 - 1) \end{bmatrix} + \frac{1}{2} \begin{bmatrix} (x_1 - 1) \\ (x_2 - 1) \end{bmatrix}^T \begin{bmatrix} 18 & 9 \\ 9 & 0 \end{bmatrix} \begin{bmatrix} (x_1 - 1) \\ (x_2 - 1) \end{bmatrix} \tag{b}$$

where $f(\mathbf{x}^*) = 3$ has been used. Simplifying the expression by expanding vector and matrix products, we obtain Taylor's expansion for $f(\mathbf{x})$ about the point (1, 1) as

$$\bar{f}(\mathbf{x}) = 9x_1^2 + 9x_1 x_2 - 18x_1 - 6x_2 + 9 \tag{c}$$

EXAMPLE 4.9 LINEAR TAYLOR'S EXPANSION OF A FUNCTION

Obtain a linear Taylor's expansion for the function

$$f(\mathbf{x}) = x_1^2 + x_2^2 - 4x_1 - 2x_2 + 4 \tag{a}$$

at the point $\mathbf{x}^* = (1, 2)$. Compare the approximate function with the original function in a neighborhood of the point (1, 2).

Solution

The gradient of the function at the point (1, 2) is given as

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} (2x_1 - 4) \\ (2x_2 - 2) \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} \tag{b}$$

Since $f(1, 2) = 1$, Eq. (4.13) gives a linear Taylor's approximation for $f(\mathbf{x})$ as

$$\bar{f}(\mathbf{x}) = 1 + [-2 \;\; 2] \begin{bmatrix} (x_1 - 1) \\ (x_2 - 2) \end{bmatrix} = -2x_1 + 2x_2 - 1 \tag{c}$$

To see how accurately $\bar{f}(x)$ approximates the original $f(x)$ in the neighborhood of (1, 2), we calculate the functions at the point (1.1, 2.2), a 10% change in the point as $\bar{f}(x) = 1.20$ and $f(x) = 1.25$. We see that the approximate function underestimates the real function by 4%. An error of this magnitude is quite acceptable in many applications. However, note that the errors will be different for different functions and can be larger for highly nonlinear functions.

Taylor Series Expansion

Assume *f*(x) is:

1. Continuous function of a single variable x
2. Differentiable n times
3. x ∈ *S*, where *S* is non-empty, closed, and bounded
4. therefore x* is a possible optima

$$f(x) = f(x^*) + \frac{df(x^*)}{dx}(x - x^*) + \frac{1}{2}\frac{d^2 f(x^*)}{dx^2}(x - x^*)^2 + \ldots$$
$$\frac{1}{3!}\frac{d^3 f(x^*)}{dx^3}(x - x^*)^3 + \frac{1}{4!}\frac{d^4 f(x^*)}{dx^4}(x - x^*)^4 + h.o.t.$$

Taylor Series Approximations

let $x - x^* = d$, a small change or vicinity of x^*...then

$$f(x) = f(x^* + d)$$

$$f(x) = f(x^*) + \frac{df(x^*)}{dx} d + \frac{1}{2} \frac{d^2 f(x^*)}{dx^2} d^2 + \frac{1}{3!} \frac{d^3 f(x^*)}{dx^3} d^3 + h.o.t.$$

$$f(x) \approx f(x^*) + \frac{df(x^*)}{dx} d \qquad \text{First Order Approximation}$$

$$f(x) \approx f(x^*) + \frac{df(x^*)}{dx} d + \frac{1}{2} \frac{d^2 f(x^*)}{dx^2} d^2 \qquad \text{Second Order Approximation}$$

Single variable minimization

Given that x* is the minimum of *f*(x), then any movement away from x* is “uphill”, therefore to guarantee that a move goes uphill

$$\Delta f(x) \geq 0$$

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

$$\frac{1}{3!} f'''(x^*)d^3 + \frac{1}{4!} f^{iv}(x^*)d^4 + h.o.t.$$

First-order necessary condition

$$f'(x^*) = 0$$

Stationary point=max,min,neither

Any points satisfying $f'(x^*) = 0$

Are called “stationary” points.

Those points are a:

1. Min pt, or
2. Max pt, or
3. Neither (i.e. an inflection pt)

We need another test!

Second-order sufficient condition

Look at second order term

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

$$\frac{1}{3!} f'''(x^*)d^3 + \frac{1}{4!} f^{iv}(x^*)d^4 + h.o.t.$$

Second-order sufficient condition for a minimum

$$f''(x^*) > 0$$

Second-order condition?

What if $f''(x^*) < 0$

Then *f*(x*) is not a minimum of *f*(x*).

It is a maximum of *f*(x*).

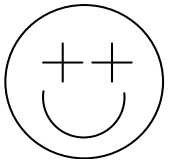
Single variable optimization

First-order necessary condition

$$f'(x^*) = 0$$

$$f''(x^*) > 0$$

Second-order sufficient condition for a minimum



$$f''(x^*) < 0$$

Second-order sufficient condition for a maximum

Higher—order tests?

What if $f''(x^*) = 0$

Then the second order test fails. We need higher order derivatives…

Min $f^{even}(x^*) > 0$

Max $f^{even}(x^*) < 0$

Second-order necessary conditions

Note the "= 0" possibility

$f''(x^*) \geq 0$

A pt not satisfying this test is not a min!

$f''(x^*) \leq 0$

A pt not satisfying this test is not a max!

Multiple variable optimization

If x* is the minimum of $f(\mathbf{x})$, then any movement away from $\mathbf{x}^*$ is "uphill".

$$\Delta f = f(\mathbf{x}) - f(\mathbf{x}^*) > 0$$

$$\Delta f = f(\mathbf{x}^* + \mathbf{d}) - f(\mathbf{x}^*) = \nabla f^T(\mathbf{x}^*)\mathbf{d} + \frac{1}{2}\mathbf{d}^T\mathbf{H}\,\mathbf{d} + R$$

$$\Delta f = \nabla f^T(\mathbf{x}^*)\mathbf{d} + \frac{1}{2}\mathbf{d}^T\mathbf{H}\,\mathbf{d} \quad \text{(second order)}$$

dot product → scalar

triple product → scalar

$$\Delta f = scalar$$

How can we guarantee that for a move in any **d,** we make away from $\mathbf{x}^*$, we go "uphill"?

First-order necessary Condition

For **x*** to be a local minimum: $\Delta f = f(\mathbf{x}) - f(\mathbf{x}^*) > 0$

$$\Delta f = \nabla f^T(\mathbf{x}^*)\mathbf{d} + \frac{1}{2}\mathbf{d}^T\mathbf{H}\,\mathbf{d}$$

1rst order term

$$\nabla f^T(\mathbf{x}^*) = \mathbf{0} \qquad \nabla f(\mathbf{x}^*) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}_{\mathbf{x}^*} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \\ 0 \end{bmatrix} = \mathbf{0}$$

Second-order sufficient condition

For **x*** to be local minimum:

$$\mathbf{d}^T \mathbf{H}\, \mathbf{d} > 0$$

That is **H(x***) must be positive definite

Remember that $\mathbf{d}^T \mathbf{H}\, \mathbf{d}$

has “quadratic form”

Quadratic form of a matrix

Given $F(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\,\mathbf{x}$ where $\mathbf{A}$ is symmetric, for example :

$$\mathbf{x}^T = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\mathbf{x}^T\mathbf{A}\,\mathbf{x} = \mathbf{x}^T \bullet (\mathbf{A}\mathbf{x}) \qquad \mathbf{A}\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\mathbf{x}^T\mathbf{A}\,\mathbf{x} = \mathbf{x}^T \bullet (\mathbf{A}\mathbf{x}) = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1^2 + x_2^2 + x_3^2$$

$$F(\mathbf{x}) = \frac{1}{2}\left(x_1^2 + x_2^2 + x_3^2\right)$$

EXAMPLE 4.10 MATRIX OF THE QUADRATIC FORM

Identify a matrix associated with the quadratic form:

$$F(x_1, x_2, x_3) = 2x_1^2 + 2x_1x_2 + 4x_1x_3 - 6x_2^2 - 4x_2x_3 + 5x_3^2 \tag{a}$$

Solution

Writing F in the matrix form ($F(\mathbf{x}) = \mathbf{x}^T\mathbf{P}\mathbf{x}$), we obtain

$$F(\mathbf{x}) = [x_1 \; x_2 \; x_3]\begin{bmatrix} 2 & 2 & 4 \\ 0 & -6 & -4 \\ 0 & 0 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{b}$$

For example, the following matrices are also associated with the same quadratic form:

$$\mathbf{P} = \begin{bmatrix} 2 & 0.5 & 1 \\ 1.5 & -6 & -6 \\ 3 & 2 & 5 \end{bmatrix}; \quad \mathbf{P} = \begin{bmatrix} 2 & 4 & 5 \\ -2 & -6 & 4 \\ -1 & -8 & 5 \end{bmatrix} \tag{c}$$

EXAMPLE 4.10 MATRIX OF THE QUADRATIC FORM

Identify a matrix associated with the quadratic form:

$$F(x_1, x_2, x_3) = 2x_1^2 + 2x_1x_2 + 4x_1x_3 - 6x_2^2 - 4x_2x_3 + 5x_3^2 \qquad \text{(a)}$$

Dividing the coefficients equally between the off-diagonal terms, we obtain the symmetric matrix associated with the quadratic form in E q. (a) as

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 2 \\ 1 & -6 & -2 \\ 2 & -2 & 5 \end{bmatrix} \qquad \text{(d)}$$

The diagonal elements of the symmetric matrix **A** are obtained from the coefficient of x_i^2 as before. The off-diagonal elements are obtained by dividing the coefficient of the term x_ix_j equally between a_{ij} and a_{ji}. Any of the matrices in Eqs. (b)–(d) gives a matrix associated with the quadratic form.