

MATH 661 Optimization

Monday 16 Sept.

- Assignment #2 due Wednesday 18 September
- Assignment #3 due Monday 23 September
- today: make sure you have version 2

Special, short asynchronous

lecture on Taylor series

in n variables (section 2.6)

Taylor series (etc.) in 1 variable

given: function $f(x)$ and basepoint $x_0 \in \mathbb{R} \rightarrow \mathbb{R}$

(f needs to be defined and differentiable
on an interval around x_0)

Taylor series: $f(x_0 + p) = f(x_0) + f'(x_0)p + \frac{1}{2}f''(x_0)p^2 + \dots + \frac{1}{k!}f^{(k)}(x_0)p^k + \dots$

kth Taylor polynomial: $q_k(p) = f(x_0) + f'(x_0)p + \frac{1}{2}f''(x_0)p^2 + \dots + \frac{1}{k!}f^{(k)}(x_0)p^k$

[casual: $f(x_0 + p) \approx f(x_0) + f'(x_0)p + \dots + \frac{1}{k!}f^{(k)}(x_0)p^k$]

Ex: find Taylor series and 4th Taylor polynomial for $f(x) = \ln x$ using base point $x_0 = 1$

Soln: $f(x) = \ln x$

$$f'(x) = \frac{1}{x} = x^{-1}$$

$$f''(x) = -x^{-2}$$

$$f'''(x) = +2x^{-3}$$

$$f^{(4)}(x) = -3 \cdot 2x^{-4}$$

$$f^{(5)}(x) = +4 \cdot 3 \cdot 2x^{-5}$$

:

$$\underline{k \geq 1:} \quad f^{(k)}(x) = (-1)^{k-1} (k-1)! x^{-k}$$

$$\begin{aligned} f(1+p) &= 0 + 1 \cdot p + \frac{1}{2} \cdot (-1) p^2 \\ &\quad + \frac{1}{3!} (+1) 2 p^3 \\ &\quad + \frac{1}{4!} (-1) 3! p^4 \\ &\quad + \frac{1}{5!} (+1) 4! p^5 \\ &\quad + \dots \\ &= p - \frac{p^2}{2} + \frac{p^3}{3} - \frac{p^4}{4} + \frac{p^5}{5} \\ &\quad - \dots (-1)^{k-1} \frac{p^k}{k} + \dots \end{aligned}$$

So:

$$g_4(p) = p - \frac{p^2}{2} + \frac{p^3}{3} - \frac{p^4}{4}$$

ideas: ① the Taylor series may only converge if p is not too large ("radius of convergence")

② the Taylor series may not equal the original function

Ex of ①: $\ln(1+p) = \sum_{j=1}^{\infty} (-1)^{j-1} \frac{p^j}{j}$

has $R=1$
radius
(from last slide)

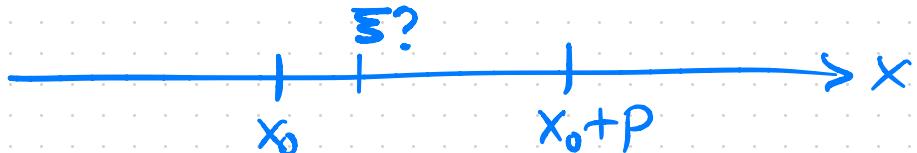
Ex of ②: $f(x) = \begin{cases} 0, & x=0 \\ e^{-1/x^2}, & x \neq 0 \end{cases}$ $x_0=0$ { don't worry about it here}

Then the series is $0 + 0p + 0p^2 + 0p^3 + \dots \neq f(p)$

- observe where I used " $=$ " versus " \approx " on previous slides ... be careful this way please!
- sometimes you want Taylor's theorem:

$$f(x_0 + p) \underset{\text{yes, } = \text{ here}}{\equiv} f(x_0) + f'(x_0)p + \dots + \frac{1}{k!} f^{(k)}(x_0) p^k + \underbrace{\frac{1}{(k+1)!} f^{(k+1)}(\xi) p^{k+1}}$$

where ξ is some number between x_0 and $x_0 + p$



- in optimization the uses of Taylor stuff are often just the linear or quadratic approximations:

$$f(x_0 + p) \approx \underbrace{f(x_0) + f'(x_0)p}_{= g_1(p)} \quad \text{linear approx.}$$

$$f(x_0 + p) \approx \underbrace{f(x_0) + f'(x_0)p + \frac{1}{2}f''(x_0)p^2}_{g_2(p)} \quad \text{quadratic approx.}$$

- but we want these in \mathbb{R}^n !

Taylor series in n variables, but only out
to 2nd order

$S \subset \mathbb{R}^n$ open

$f: S \rightarrow \mathbb{R}$ is continuous, and all its derivatives are continuous

$x_0 \in S$

then for $p \in \mathbb{R}^n$:

$$f(x_0 + p) = f(x_0) + \nabla f(x_0)^T p + \frac{1}{2} p^T \nabla^2 f(x_0) p$$

+ ...
in n variables it is not
clear (initially) what this is

We don't care about higher order:

all our optimization uses of Taylor in n variables will be

$$f(x_0 + p) \approx f(x_0) + \nabla f(x_0)^T p$$

linear approx.

or

$$f(x_0 + p) \approx f(x_0) + \nabla f(x_0)^T p + \frac{1}{2} p^T \nabla^2 f(x_0) p$$

quadratic approx

Hessian

Ex: Find the first 3 terms of the Taylor series for

$$f(x) = 3x_1^4 - x_1x_2 + 5x_1x_2^2 + 2$$

at the point $x_0 = (1, 1)^T$.

Evaluate the series for $p = (0.1, -0.1)^T$ and compare with the value of $f(x_0 + p)$.

asked
just
like
Exercise
6.4

Solution: $f(x) = 3x_1^4 - x_1x_2 + 5x_1x_2^2 + 2 \quad n=2$

$$\nabla f(x) = \begin{bmatrix} 12x_1^3 - x_2 \\ -x_1 + 10x_1x_2 \end{bmatrix}$$

$$\nabla^2 f(x) = \begin{bmatrix} 36x_1^2 & | & -1 \\ -1 & | & 10x_1 \end{bmatrix}$$

$x_0 = (1, 1)^T$ so:

$$f(x_0)$$

$$\nabla f(x_0)^T P$$

$$\frac{1}{2} p^T \nabla^2 f(x_0) p$$

$$f(x_0 + p) \approx 9 + [1 \ 1] \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} + \frac{1}{2} [1 \ 1] \begin{bmatrix} 36 & -1 \\ -1 & 10 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

$$= 9 + 11p_1 + 9p_2 + \frac{1}{2} [p_1 \ p_2] \begin{bmatrix} 36p_1 - p_2 \\ -p_1 + 10p_2 \end{bmatrix}$$

$$\begin{aligned}
 &= 9 + 11p_1 + 9p_2 + \frac{1}{2} [p_1 \ p_2] \begin{bmatrix} 36p_1 - p_2 \\ -p_1 + 10p_2 \end{bmatrix} \\
 &= 9 + 11p_1 + 9p_2 + \frac{1}{2}(36p_1^2 - p_1p_2 - p_1p_2 + 10p_2^2) \\
 &= 9 + 11p_1 + 9p_2 + 18p_1^2 - p_1p_2 + 5p_2^2
 \end{aligned}$$

and with $p = (0.1, -0.1)^T$:

$$\begin{aligned}
 q_2(p) &= \boxed{} \\
 f(x_0 + p) &= \boxed{}
 \end{aligned}$$

quadratic in $p \in \mathbb{R}^2$

I owe you these
... see Announcements!

Our main use of Taylor ideas:

- Consider a smooth, unconstrained, generally not convex optimization problem

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

- suppose $x_j \in \mathbb{R}^n$ is our current iterate in some optimization algorithm

• Then

$$g(p) = f(x_j) + \nabla f(x_j)^T p$$

is our "linear model" of f near x_j "

and

$$h(p) = f(x_j) + \nabla f(x_j)^T p + \frac{1}{2} p^T \nabla^2 f(x_j) p$$

is our "quadratic model" of f near x_j "

- the next step in the optimization algorithm is

$$x_{j+1} = x_j + p^*$$

compute this as

$$\min_p g(p)$$

$$\text{or } \min_p h(p)$$

possibly
subject
to "not too
far"
constraints