CHAPTER 3: HIGHER ORDER DERIVS: MAXIMA & MINIMA

One variable:

↳ sign of first deriv says if incr or decreasing

↳ extrema (max/min) at endpoints or where first deriv is zero (critical points).

↳ tell which by second deriv test

TAYLOR'S THM

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \frac{f'''(x_0)}{3!}(x-x_0)^3 + \dots$$

↳ truncate and control the error

↳ how to generalize, as get matrices and vectors

$f'(x_0)(x-x_0)$ $\Longrightarrow$ becomes $\Longrightarrow$ $\nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$

↳ what about $f''(x_0)(x-x_0)^2$?

$(\vec{x}-\vec{x}_0)^T\, Hf(\vec{x}_0)\, (\vec{x}-\vec{x}_0)$ with Hf a matrix of second derivs

↳ utility of Taylor: replace complicated fn with simpler

Proof: Integral version in book, will sketch another here

Claim $f(x) = \sum_{i=1}^{n} \frac{f^{(i)}(x_0)}{i!}(x-x_0)^i + R_{n+1}(x_0, x)$

small rel to $(x-x_0)^n$

↳ MVT: $\frac{f(x)-f(x_0)}{x-x_0} = f'(c_1)$

$$\Rightarrow f(x) = f(x_0) + f'(c_1)(x-x_0)$$

but $\frac{f'(c_1)-f'(x_0)}{c_1-x_0} = f''(c_2)$

$\Rightarrow f'(c_1) = f'(x_0) + f''(c_2)(c_1-x_0)$

$$= f(x_0) + f'(x_0)(x-x_0) + f''(c_2)(c_2-x_0)(x-x_0)$$

CHAPTER 3 (CONT)

For many purposes linear approx suffices.

Common expansions:

$$e^{x} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \frac{e^{ix} - e^{-ix}}{2}$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots = \frac{e^{ix} - e^{-ix}}{2i}$$

$$\cos x + i \sin x = e^{ix}$$

↳ Theorem that $e^{a} e^{b} = e^{a+b}$

↳ can derive all trig identities!

$$1 = e^{ix} e^{-ix} = (\cos x + i \sin x)(\cos x - i \sin x) = \cos^2 x + \sin^2 x$$

$$e^{i(\theta + \varphi)} = e^{i\theta} e^{i\varphi}$$

↳ $\cos(\theta+\varphi) + i \sin(\theta + \varphi) = (\cos\theta + i \sin\theta)(\cos\varphi + i \sin\varphi)$

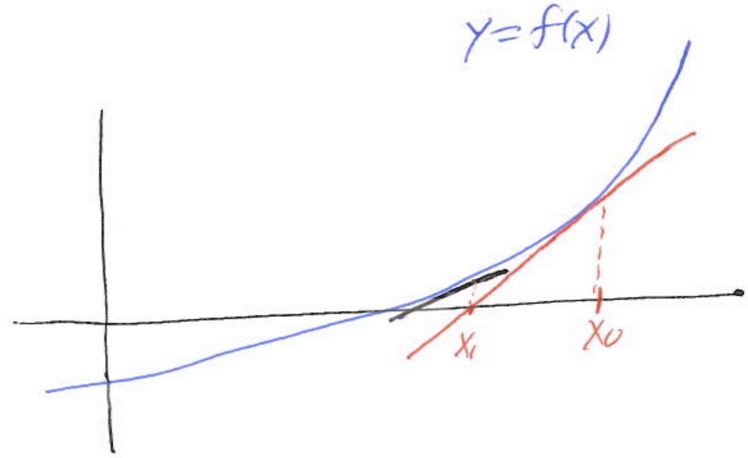
$= (\cos\theta \cos\varphi - \sin\theta \sin\varphi) + i(\sin\theta \cos\varphi + \cos\theta \sin\varphi)$

$$\log(1-x) = -\left(x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots\right)$$

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$$

CHAPTER 3: (CONT)

NEWTON'S METHOD



Find root of $f(x)=0$
Assume f is "nice" and differentiable

Step 1: Guess x_0 for root. If $f(x_0)=0$ done else continue

Step 2: Consider point $(x_0, f(x_0))$ on graph. Tangent line has slope $f'(x_0)$. Approx fn by tangent line, see where it crosses the x-axis, call that x_1.

↳ tangent line: $y - f(x_0) = f'(x_0)(x-x_0)$

intercept: $0 - f(x_0) = f'(x_0)(x_1 - x_0) \to x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$

Lather, rinse, repeat: incredible fast

↳ Numerous applications

↳ Fractals

If $f(x)=x^2-3$ then $x_{n+1} = x_n - \frac{x_n^2-3}{2x_n} = \frac{1}{2}\left(x_n + \frac{3}{x_n}\right)$

↳ sequence converges really fast!

Much better than Divide and Conquer: Halve error each iteration

SECTION 3.1: ITERATED PARTIAL DERIVS

GOAL: When does $\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)$ equal $\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)$?

Notation: $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)$, $\frac{\partial^2 f}{\partial x_i^2} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_i}\right)$

Computing easy: $f(x) = x^2 y \sin x$ Then

$$\partial f/\partial x = 2xy\sin x + x^2 y\cos x$$

$$\partial f/\partial y = x^2 \sin x$$

$$\partial^2 f/\partial x\partial y = 2x\sin x + x^2\cos x = \partial^2 f/\partial y\partial x$$

Defn: C^2: f is of class C^2 if all partial derivs $\partial f/\partial x_i$ exist and further each of these have continuous partial derivatives.

THM: EQUALITY OF MIXED PARTIAL DERIVS

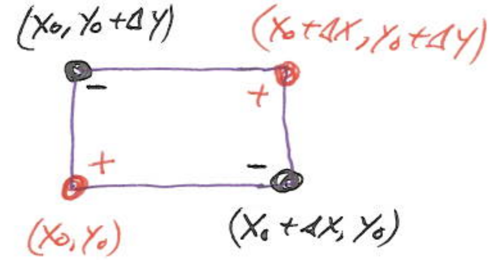
If f is of class C^2 then $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$

↳ Not always equal

↳ Exercise 24: $f(x,y) = \begin{cases} xy(x^2-y^2)/(x^2+y^2) & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$

↳ Show $\frac{\partial^2 f}{\partial x \partial y}(0,0) = 1$ but $\frac{\partial^2 f}{\partial y \partial x}(0,0) = -1$

Proof of Equality of Mixed Partials Thm



$$S(\Delta x,\Delta y) = f(x_0+\Delta x, y_0+\Delta y) - f(x_0+\Delta x, y_0) - f(x_0, y_0+\Delta y) + f(x_0, y_0)$$

Holding $y_0, \Delta y$ fixed: $g(x) := f(x, y_0+\Delta y) - f(x, y_0)$

$$\Longrightarrow S(\Delta x, \Delta y) = g(x_0+\Delta x) - g(x_0)$$

$$= g'(\tilde{x})\Delta x \quad \tilde{x} \text{ b/w } x_0 \text{ and } x_0+\Delta x \text{ (MVT)}$$

$$= \left[\frac{\partial f}{\partial x}(\tilde{x}, y_0+\Delta y) - \frac{\partial f}{\partial x}(\tilde{x}, y_0)\right]\Delta x$$

$$= \frac{\partial^2 f}{\partial y\,\partial x}(\tilde{x},\tilde{y})\,\Delta x\,\Delta y \quad \text{by MVT}$$

As $\partial^2 f/\partial y\partial x$ is cont

$$\frac{\partial^2 f}{\partial y\,\partial x}(x_0, y_0) = \lim_{(\Delta x,\Delta y)\to(0,0)} \frac{S(\Delta x,\Delta y)}{\Delta x\,\Delta y}$$

As $S(\Delta x, \Delta y)$ is symmetric, similar calculation shows above is also $\partial^2 f/\partial x\partial y$, completing the proof. ∎

See book for many famous partial diff eqs

Homework: #1, #11

Suggested: #16, #19

Section 3.2: Taylor's Thm

Question: When does $f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots$?

Danger: $f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$ This function is

infinitely diff. By L'Hopital all derivs vanish at the origin,

and thus only equals its Taylor series at $x=0$ (unimpressive).

Advanced: Alternate proof of Taylor

$f_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k$, $h(x) = f(x) - f_n(x)$, $h(0) = 0$

$$\begin{aligned}
h(x) &= h(x) - h(0) \quad \text{(Adding zero, powerful technique)} \\
&= h'(c_1)(x-0) \quad \text{with } c_1 \in [0,x] \\
&= \left(f'(c_1) - f_n'(c_1)\right) x \\
&= \left(f'(c_1) - \sum_{k=1}^{n} \frac{f^{(k)}(0)}{(k-1)!} c_1^{k-1}\right) x \\
&= h_1(c_1)\, x \quad \text{with } h_1(\xi) = f'(\xi) - \sum_{k=1}^{n} \frac{f^{(k)}(0)}{(k-1)!} \xi^{k-1}
\end{aligned}$$

Apply MVT to h_1, noting $h_1(0) = 0$

$$\begin{aligned}
h_1(c_1) = h_1(c_1) - h_1(0) &= h_1'(c_2)\, c_1 \quad , \quad c_2 \in [0, c_1] \subset [0,x] \\
&= \left(f''(c_2) - \sum_{k=2}^{n} \frac{f^{(k)}(0)}{(k-2)!} c_2^{k-2}\right) c_1 = h_2(c_2)\, c_1
\end{aligned}$$

⋮ continue

$$h(x) = f^{(n+1)}(c_{n+1}) \cdot c_{n-1} c_{n-2} \cdots c_2 c_1 x$$

↳ Thus $|h(x)| = |f(x) - f_n(x)| \leq \max_{c \in [0,x]} |f^{(n+1)}(c)| \cdot |x|^{n+1}$

Hessian: $\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j} \right) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$

Also denote as $Hf(\vec{x})$

TAYLOR'S THM (SEVERAL VARIABLES)

$$\boxed{f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot \vec{h} + \frac{1}{2} \vec{h}^T Hf(\vec{x}_0) \vec{h} + \cdots}$$

Example: Find second order Taylor series of $f(x,y) = \sin(x+2y)$

↳ Means get all terms that are constant, one or two derivs

↳ book does by brute force, faster way

↳ $\sin\theta = \theta - \dfrac{\theta^3}{3!} + \cdots$

Thus $\sin(x+2y) = x+2y - \dfrac{(x+2y)^3}{3!} + \cdots$

↳ Hence second order expansion is just $x+2y$

Homework: #2, #3

Suggested: #7

Section 3.3: Extrema of Real Valued Fns

Def'n: local extremum: f has a local max (resp, local min) at $\vec{x}_0$ if there is a nbhood of $\vec{x}_0$ such that for all $\vec{x}$ in this nbhood, $f(\vec{x}) \leq f(\vec{x}_0)$ (resp $f(\vec{x}) \geq f(\vec{x}_0)$).

Thm: First Deriv Test for Local Extrema

f: open $U \subset \mathbb{R}^n \to \mathbb{R}$. If f is diff and $\vec{x}_0 \in U$ is a local extremum, then $Df(\vec{x}_0) = \vec{0}$ (ie, $\vec{x}_0$ is a critical point)

Proof: Assume f has local max

Let $c(t) = \vec{x}_0 + t\vec{h}$

Then $g(t) = f(c(t))$ has local max at $t=0$

↳ Chain rule: $g'(0) = Df(\vec{x}_0)\,\vec{h} = 0$

↳ as true for all $\vec{h}$, have $Df(\vec{x}_0) = \vec{0}$ ∎

- Note we reduced to one-dimension!

- There is a generalized second deriv test to see if max or min, but without linear algebra it is hard to see what it says.

Next few pages advanced material on max/min

Homework: #7, #22, #36

Suggested: #18, #23, #28, #41

3.3. EXTREMA OF REAL VALUED FNS (ADVANCED)

Thm: f cont fn on finite interval $[a,b]$. Then f attains max/min.

↳ needed result: $\{x_n\} \subset [a,b] \Rightarrow$ subseq converges

Proof: wlog, $[a,b] = [0,1] = I_0$

As ∞ many terms, either $[0,1/2]$ or $[1/2,1]$
contains ∞ many, call I_1, divide
Choose x_{n_1} any element in I_1, study seq from n_1 on

Note $|x_{n_m} - x_{n_{m'}}| \le 2^{-N}$ if $n_m, n_{m'} \ge N$

Dist b/w left and right endpoints $\to 0$

needed result: f bounded

Proof: assume not. Let x_n ($\ne x_m$ for any $m<n$) be st $f(x_n) \ge n$.

Subseq converges, limit point $\tilde{x}$, $f(\tilde{x})$ must be ∞

By continuity, f is ∞ in nbhood $\tilde{x}$

needed defn: sup: supremum of a seq of values is a number c st
every value is at most c, and if $c' < c$ then some value
exceeds c'. Note unique, inf defined

Main Result: f attains max

Proof: f bounded, $c = \sup_{x \in [a,b]} f(x)$. Assume not

Choose x_n st $f(x_n) > f(x_{n-1})$ and $c - f(x_n) < \frac{1}{n}$ (defn sup)

↳ choose subseq st $f(x_{n_k}) \to c$

↳ choose subsubseq st $x_{n_{k_\ell}} \to \tilde{x}$

Claim: $f(\tilde{x}) = c$

↳ if not, choose $\varepsilon < \frac{c - f(\tilde{x})}{2}$

By cont, $|x - \tilde{x}| < \delta \to |f(x) - f(\tilde{x})| < \varepsilon$

Violates assumps on seq $x_{n_{k_\ell}}$

Like Divide + Conquer vs Newton's Method: more assume, better results

Ask about Newton's Method

3.3. Extrema of Real Valued Fns

(ADVANCED)

Defn local max/min: nbhood V of x_0 st $f(x) \lesseqgtr f(x_0)$ $\forall x \in V$

Thm: 1st Deriv Test: f: open $U \to \mathbb{R}$ diff, $x_0 \in U$ local extremum

Then $Df(\vec{x}_0) = \vec{0}$ ($\vec{x}_0$ is a critical pnt of f)

Proof: ① could use Taylor Thm w/ error (if C^2)

② $g(t) = f(\vec{x}_0 + t\vec{v})$

one-var: $g'(t) = 0 \Rightarrow Df(\vec{x}_0)\vec{v} = 0$

take $\vec{v} = Df(\vec{x}_0)$, so $\vec{0}$

Lin Alg Digression

A $n \times n$ matrix, A is pos-def if $\forall \vec{x} \neq \vec{0}$, $\vec{x}^T A \vec{x} > 0$

Note $\vec{x}^T A \vec{x} = \sum a_{ij} x_i x_j$ wlog assume $A = A^T$ real symm

Lemma: A pos def $\to \exists M$ st $\vec{x}^T A \vec{x} \geq M \|\vec{x}\|^2$

Proof: wlog $\|\vec{x}\| = 1$ (trivial if $\|\vec{x}\| = 0$)

let $M = \min_{\|\vec{x}\|=1} \vec{x}^T A \vec{x}$: poly

↳ showed cont fns attain min/max,

min > 0 as pos def

Eigenvalues/vectors: $\vec{v} \neq \vec{0}$: $A\vec{v} = \lambda \vec{v}$ (same dir, different magnitude)

↳ means $(A - \lambda I)\vec{v} = \vec{0}$: if matrix invertible then $\vec{v} = \vec{0}$

so solving $\mathrm{Det}(A - \lambda I) = 0$: Fund Thm Alg

$A = A^T \to$ all $\lambda \in \mathbb{R}$ and $\exists Q$ st $A = Q^{-1}\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} Q$, $Q^{-1}Q = I = QQ^{-1}$

and $Q^{-1} = Q^T$!

So $\vec{x}^T A \vec{x} = \vec{x}^T Q^T \Lambda Q \vec{x} = \vec{y}^T \Lambda \vec{y}$

Easy to tell if pos def: each $\lambda_i > 0$!

Thm: 2nd Deriv Test: $f: U \to \mathbb{R}$ is $\mathcal{C}^3$, $\vec{x}_0$ critical point, Then max(min) if $Hf(\vec{x}_0)$ is pos (neg) def

Proof: $f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \underbrace{Df(\vec{x}_0) \cdot \vec{h}}_{0} + \vec{h}^T Hf(\vec{x}_0) \vec{h} + R_2(\vec{x}_0, \vec{h})$

with $R_2(\vec{x}_0, \vec{h}) / \|h\|^2 \to 0$

for $\|\vec{h}\|$ suff small, sign of $f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0)$ controlled by Hessian

Read book for pos def tests

↳ make more sense after Lin Alg

HW: pg 222: #7, #22, #36

Suggested: #18, #23, #28, #41

Section 3.4: Constrained Extrema + Lagrange Multipliers

Subtle, important point: one thing to find candidates for max/min; another to prove that the function attains a max/min. These proofs are sketched in the previous pages.

Method of Lagrange Multipliers

$f, g: U \subset \mathbb{R}^n \to \mathbb{R}$ are C^1, $\vec{x}_0 \in U$ st $g(\vec{x}_0) = c$, $S = \{\vec{x} \in U : g(\vec{x}) = c\}$. Assume $\nabla g(\vec{x}_0) \neq \vec{0}$. Let $f|_S$ be f restricted to S. Then $f|_S$ has extremum at $\vec{x}_0$ if and only if there is a λ st $\nabla f(\vec{x}_0) = \lambda \nabla g(\vec{x}_0)$.

Proof: S level set: $\frac{d}{dt}\left(g(c(t))\right) = \nabla g(\vec{x}_0) \cdot c'(0) = 0$

where $c'(0)$ is any vector tangent to S at $\vec{x}_0$.

↳ If f has max/min at $\vec{x}_0$ then $\frac{d}{dt}\left(f(c(t))\right) = \nabla f(\vec{x}_0) \cdot c'(0) = 0$

↳ Thus $\nabla f(\vec{x}_0)$ and $\nabla g(\vec{x}_0)$ perpendicular to all tangent dirs. Only one dir left, so $\nabla f(\vec{x}_0)$ and $\nabla g(\vec{x}_0)$ in that dir, and hence parallel! ∎

Interpretation: $\nabla g(\vec{x}_0)$ is normal to surface, says max/min means $\nabla f(\vec{x}_0)$ normal to surface: if not, flow in proper direction and increase.

Section 3.4 (cont)

Application: $\frac{\partial f}{\partial x_i}(\vec{x_0}) = \lambda \frac{\partial g}{\partial x_i}(\vec{x_0})$ and $g(\vec{x_0}) = c$

↳ have $n+1$ equations in $n+1$ variables: should be solvable

Example: $f(x,y) = 3x+2y$ $\quad g(x,y) = 2x^2+3y^2 = 3$ (ellipse)

$\nabla f(x,y) = (3,2)$ $\quad \nabla g(x,y) = (4x, 6y)$

↳ $(3,2) = \lambda(4x, 6y)$ and $2x^2+3y^2 = 3$

↳ $\left.\begin{array}{l} 4\lambda x = 3 \\ 6\lambda y = 2 \end{array}\right\} \xrightarrow{\text{ratio}} \frac{4x}{6y} = \frac{3}{2}$ or $y = \frac{4}{9}x$

$2x^2+3y^2=3$

↳ Thus $2x^2 + 3\left(\frac{4}{9}x\right)^2 = 3$

So $2x^2 + \frac{16}{27}x^2 = 3 \Longrightarrow 18x^2 = 81$

↳ Thus $x = \pm \frac{3\sqrt{2}}{2}$, $y = \pm \frac{2\sqrt{2}}{3}$

Have four candidate points: check

Note: will not do Section 3.5 (Inverse + Implicit Fn Thms). But, important results – worth reading.

Homework: #2, #10

Suggested: #20, #27