M&T ch 4.3 9/29/2024

2^{nd} Deriv Test for Constrained Extrema. Bordered Hessian ⓺a

We seek to classify stationary pts a.k.a. critical pts of $f|_M$ where $M := g^{-1}(c)$

$f : \mathbb{R}^n \longrightarrow \mathbb{R}$ constraint $g : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ a submfd of $\mathbb{R}^n$

M&T deal with only the 2-d case $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ subj to $g(x,y) = c$

(I will give a more general discussion soon following Robinson's online book 'Intro to Math Optimizatio and associated online Supplement 'Second Deriv Test for Constrained Extrema')

Define $h(x,y,\lambda) := f(x,y) - \lambda(g(x,y) - c)$ Robinson calls this $\mathcal{L}$

M&T's arg (which seems to only work for $\mathbb{R}^2$) is briefly:

look for critical pts of h $D_x h = 0 \Rightarrow [\, f_x - \lambda g_x \quad f_y - \lambda g_y \,] \overset{!}{=} [\, 0 \quad 0 \,]$

Reduce to a 1-d problem by $\left.\frac{\partial g}{\partial y}\right|_{x_0} \neq 0$ (assumption for Imp Fcn Thm - ch 4.4)

Then $M = \text{Graph}(\varphi)$ where $y = \varphi(x)$ in nbhd of $x_0\, y_0$

$$f(x,y) = f(x, \varphi(x))$$

apply the chain rule and by a lot of trickery we get

$$\frac{d^2 f}{dx^2} = -\frac{1}{g_y^2} \det \left[\begin{array}{c:cc} 0 & g_x & g_y \\ \hdashline g_x & h_{xx} & h_{xy} \\ g_y & h_{yx} & h_{yy} \end{array}\right]$$

The factorization tricks don't seem to apply beyond $\mathbb{R}^2$

So I'm going to abandon M&T's discussion.

Remark relevant to next page:

why define $\mathcal{L}(\vec{\lambda}, \vec{x}) := f(x) - \sum^{k} \lambda_i (g^i(x) - c^i)$? why $g - c$?

Then at a critical pt $\langle \underset{\sim}{\lambda}, \underset{\sim}{x} \rangle$ $\nabla \mathcal{L} = \vec{0}$ because $g(\underset{\sim}{x}) = c$

$g^{ⓘ}(\underset{\sim}{x}) = c^{ⓘ}$

I'm not sure what is the advantage

R. Clark Robinson
math.northwestern.edu

2^{nd} Deriv Test for Constrained Extrema

8/21/2024

$$f: \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$g: \mathbb{R}^n \longrightarrow \mathbb{R}^k \qquad k < n \qquad \text{Let } M = g^{-1}(c)$$

submfd of $\mathbb{R}^n$ (so Dg_x maps onto $\mathbb{R}^k$ $\forall\, x \in g^{-1}(c)$)

We seek to classify extrema of $f/_M$

We know the Lagrange multiplier ~~convention~~ condition is satisfied at an extrema

$$\underbrace{(\lambda, x)}$$

we write λ first to make a matrix nicer

$$\nabla f = \sum_{i=1}^{k} \lambda_i \nabla g^i_x \qquad Df_x(\cdot) = \sum \lambda_i Dg^i_x(\cdot)$$

That is $D_x\left(f - \sum \lambda_i g^i\right) = 0$

Let $\langle \underline{\lambda}, a \rangle$ be a point where Lag Mult eq is satisfied

$\mathcal{L}(\underline{\lambda}, a) = f(a)$

$\nabla \mathcal{L}_{(\underline{\lambda}, a)} =$

Define $\mathcal{L}(\lambda, x) := f(x) - \sum_{i=1}^{k} \lambda_i \left(g^i(x) - c^i\right)$

Lets write this out for an example ~~(we can also drop the c)~~

$$\mathcal{L}(\overbrace{\lambda, \mu}^{k=2}, \overbrace{x, y, z}^{n=3}) = f(x,y,z) - \lambda\left(g^{\textcircled{1}}(x,y,z) - c^{\textcircled{1}}\right) - \mu\left(g^{\textcircled{2}}(x,y,z) - c^{\textcircled{2}}\right)$$

Remark: $D\mathcal{L}_{(\underline{\lambda}\, \underline{x})} = [0\ 0\ 0\ 0\ 0]$

$$D\mathcal{L}_{(\lambda x)} = [\mathcal{L}_\lambda\ \mathcal{L}_\mu\ \mathcal{L}_x\ \mathcal{L}_y\ \mathcal{L}_z]$$

$$= [\underset{a}{(-g^1 + c^1)}\ \underset{b}{(-g^2 + c^2)}\ \underset{c}{(f_x - \lambda g^1_x - \mu g^2_x)}\ \underset{d}{(f_y - \lambda g^1_y - \mu g^2_y)}\ \underset{e}{(f_z - \lambda g^1_z - \mu g^2_z)}]$$

$$D^2\mathcal{L}_{(\lambda, x)} = \left[\begin{array}{cc|ccc} a_\lambda & a_\mu & a_x & a_y & a_z \\ b_\lambda & b_\mu & b_x & b_y & b_z \\ \hline c_\lambda & c_\mu & & & \\ d_\lambda & d_\mu & & D^2_x f - \sum \lambda_i D^2 g^i & \\ e_\lambda & e_\mu & & & \end{array}\right] = \left[\begin{array}{cc|ccc} 0 & 0 & -g^{\textcircled{1}}_x & -g^{\textcircled{1}}_y & -g^{\textcircled{1}}_z \\ 0 & 0 & -g^{\textcircled{2}}_x & -g^{\textcircled{2}}_y & -g^{\textcircled{2}}_z \\ \hline -g^{\textcircled{1}}_x & -g^{\textcircled{2}}_x & (f_{xx} - \lambda g^{\textcircled{1}}_{xx} - \mu g^{\textcircled{2}}_{xx}) & \cdot & \cdot \\ -g^{\textcircled{1}}_y & -g^{\textcircled{2}}_y & \cdot & \cdot & \cdot \\ -g^{\textcircled{1}}_z & -g^{\textcircled{2}}_z & \cdot & \cdot & \cdot \end{array}\right]$$

This is the 'bordered Hessian'
Dg_x is the "border"

It is more manageable to write out a $k=1$ $n=2$ example

$$D^2\mathcal{L}_{(\lambda, x)} = \left[\begin{array}{c|cc} 0 & -g_x & -g_y \\ \hline -g_x & (f_{xx} - \lambda g_{xx}) & (f_{xy} - \lambda g_{xy}) \\ -g_y & (f_{yx} - \lambda g_{yx}) & (f_{yy} - \lambda g_{yy}) \end{array}\right]$$

$$\begin{array}{c|c|c} & k & n \\ \hline k & 0 & -Dg_x \\ n & -(Dg_x)^T & D^2 f_x - \lambda^T D^2 g_x \end{array}$$

Evaluated at the critical pt $(\lambda^*, x^*) \rightarrow$ easier to write $\longrightarrow (\underline{\lambda}, \underline{x})$

8/29/2024

KEY RESULT

Lemma 1

$\langle \underline{\lambda}, \underline{x} \rangle$ Satisfies Lag Mult eq: $Df_{\underline{x}} = \sum_{i=1}^{m} \lambda_i Dg^{\textcircled{i}}_{\underline{x}}$

Let $M := g^{-1}(c)$

σ is a curve on M $\quad \sigma(0) = \underline{x}$

$\sigma'(0) = v$

$$\Rightarrow \frac{d^2}{dt^2} f(\sigma(t))\Big|_{t=0} = v^T\left[D^2_x \mathcal{L}\right]v \text{ at } \langle \underline{\lambda}, \underline{x} \rangle$$

Pf. $f: \mathbb{R}^n \longrightarrow \mathbb{R}$

$$\frac{d}{dt}f(\sigma(t)) = Df_{\sigma(t)}\dot\sigma(t) = [D_1 f \dots D_n f]\begin{bmatrix}\dot\sigma^{\textcircled{1}}\\ \vdots \\ \dot\sigma^{\textcircled{n}}\end{bmatrix} = \sum^{n} (D_i f)_{\sigma} \dot\sigma^i$$

$$\frac{d^2}{dt^2}f(\sigma(t)) = \frac{d}{dt}\sum_i \underbrace{(D_i f)_{\sigma(t)}}_{F_i} \dot\sigma^i = \sum_i \frac{d}{dt}\left(F_i(\sigma)\dot\sigma^i\right)$$

$$= \sum_i^n \left(\left(\sum_j^n (D_j F_i)\dot\sigma^{\textcircled{j}}\right)\dot\sigma^{\textcircled{i}} + F_i(\sigma)\ddot\sigma^{\textcircled{i}}\right)$$

$$= [\dot\sigma^{\textcircled{1}} \dots \dot\sigma^{\textcircled{n}}]\left[D^2 f_{\sigma(t)}\right]\begin{bmatrix}\dot\sigma^{\textcircled{1}}\\ \vdots \\ \dot\sigma^{\textcircled{n}}\end{bmatrix} + Df_{\sigma}(\ddot\sigma) \quad \circledast$$

Evaluate at $t=0$

$$\frac{d^2}{dt^2}f(\sigma(0)) = [v^1 \dots v^n]\left[D^2 f_{\underline{x}}\right]v + \underbrace{Df_{\underline{x}}(\ddot\sigma(0))}$$

$$= \underbrace{\sum^{m} \lambda_j Dg^{\textcircled{j}}_{\underline{x}}(\ddot\sigma)} \text{ by Lag Mult cond}$$

Step 2

Now consider $g^{\textcircled{j}}(\sigma(t)) = c^{\textcircled{j}}$

apply $\circledast$: $\quad 0 = \frac{d^2}{dt^2}\left(g^{\textcircled{j}}(\sigma(t))\right) = \dot\sigma^T\left[D^2 g^j_{\underline{x}}\right]\dot\sigma + Dg^j_{\sigma}(\ddot\sigma)$

Evaluate at $t=0$ and apply $\sum_{j=1}^{m} \lambda_j$ to both sides:

$$0 = \sum_{j=1}^{m} \lambda_j v^T\left[D^2 g^j_{\underline{x}}\right]v + \sum_{j=1}^{m} \lambda_j Dg^{\textcircled{j}}_{\underline{x}}(\ddot\sigma(0))$$

$$-\sum_{j=1}^{m} \lambda_j v^T\left[D^2 g^j_{\underline{x}}\right]v = \sum_{j=1}^{m} \lambda_j Dg^j_{\underline{x}}(\ddot\sigma(0))$$

Step 3 subs this into expression for $\frac{d^2}{dt^2}f(\sigma(0))$

$$\frac{d^2}{dt^2}f(\sigma(0)) = v^T\left[D^2 f_{\underline{x}}\right]v - \sum_{j=1}^{m} \lambda_j v^T\left[D^2 g^j_{\underline{x}}\right]v = v^T\underbrace{\left[D^2 f_{\underline{x}} - \sum_{j=1}^{m} \lambda_j D^2 g^{\textcircled{j}}_{\underline{x}}\right]}_{D^2_x \mathcal{L}}v$$

□

9/14/2024

Recall from earlier in this chapter (and better Avez DC ch 10) for unconstrained f

Avez Thm 10.4 $D^2 f_{\underline{x}}$ exists, f has local min at $\underline{x}$ [$-f$ has local max] $\Rightarrow$ $Df_{\underline{x}} = 0$, $D^2 f_{\underline{x}}(v,v) \geq 0 \quad \forall v \in \mathbb{R}^n,\ v \neq 0$ pos semi-def

Converse Thm 10.6 $Df_{\underline{x}} = 0$ critical pt, $D^2 f_x(v,v) \underset{\neq}{\geq} 0 \quad \forall v \neq 0$ $\Rightarrow$ f has strict local min at $\underline{x}$.

(There is counterexample for $D^2 f_x \geq 0$)

Now let's restate the set up

$f: \mathbb{R}^n \longrightarrow \mathbb{R}$ we consider $f|_M$ where $M = g^{-1}(c)$

Constraint $g: \mathbb{R}^n \longrightarrow \mathbb{R}^k$ $\underline{x} \in M$ and $Dg_{\underline{x}}$ maps onto $\mathbb{R}^k$ ($Dg_{\underline{x}}$ has rank k)

[so M is a submfd at least in nbhd of $\underline{x}$]

We know $T_{\underline{x}} M = \ker(Dg_{\underline{x}})$ from G&P

The pt $\langle \underline{\lambda}, \underline{x} \rangle$ satisfies Lagrange multiplier eq $Df_x = \sum^k \lambda_j Dg^{ⓙ}_{\underline{x}}$

$\mathcal{L} := f - \sum^k \lambda_j (g^j - c^j) \qquad A := D^2_x \mathcal{L} \qquad H = D^2 \mathcal{L}_{(\underline{\lambda}, \underline{x})}$

Robinson's Thm 2 (really COR to 'Key Result' Lemma)

$f|_M$ has a local max at $\underline{x}$ $\Rightarrow$ $v^T \overset{A}{[D^2_x \mathcal{L}]} v \leq 0 \quad \forall\ v \in T_{\underline{x}} M = \ker(Dg_x)$

strict local max $\Rightarrow$ $v^T A v < 0$

$f|_M$ has local min at $\underline{x}$ $\Rightarrow$ $v^T A v \geq 0$

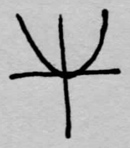
strict local min $\Rightarrow$ $v^T A v > 0$

Pf $\left.\frac{d^2}{dt^2} f(\sigma(t))\right|_{t=0} = (\dot{\sigma}(0))^T A \dot{\sigma}(0)$ from the Lemma and $v = \dot{\sigma}(0) \in T_{\underline{x}} M$

$\underline{x} = \sigma(0)$

From 1 variable calc $(f \circ \sigma): \overset{\mathbb{R}}{I} \longrightarrow \mathbb{R}$ has a min at $t=0$ $\Rightarrow f''(0) \geq 0$ concave up

$\Rightarrow v^T A v \geq 0$ and since this holds for arb σ $\sigma(0) = \underline{x}$ we get any $v \in T_{\underline{x}} M$



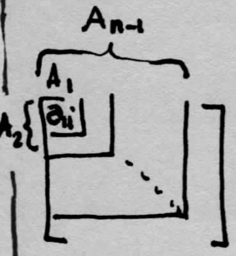
if σ is this curve along bottom, it is not an isolated min. $f''(0) = 0$

$\square$

But our purpose is to consider the <u>converse</u>: A being pos def giving $\underline{x}$ being a min. But A must be restricted to $T_{\underline{x}}M$ <u>not</u> $\mathbb{R}^n$ so usual techniques to show matrix A is pos def won't work.

ASIDE: The usual technique would be Sylvester's det thm

If all principle submatrices, starting from upper left corner, have pos dets, then A is pos def



Lets prove this, following Strang LAAIA ch 6 write-up sheets ②-③

(1) we know from Strang ch 4.4 that $\det A_k = d_1 d_2 \cdots d_k$ product of the Gaussian elimination pivots. A priori, a pivot is non-zero and since we assume all dets in the chain are pos, $\det A = \det A_n$ pos $\Rightarrow$ A non-singular $\Rightarrow$ there are n pivots.

(2) $$d_k \equiv \frac{(d_1 \cdot d_2 \cdots d_k)}{(d_1 \cdots d_{k-1})} = \frac{\det A_k > 0}{\det A_{k-1} > 0} \Rightarrow d_k \text{ is pos } \forall k$$

(3) Since A is sym, we can do "Cholesky" factorization $A = LDL^T$
Where D is diag matrix of pivots
Then $$x^T A x = x^T (LDL^T) x = (x^T L) D (L^T x) = \sum_{i=1}^{n} d_i \left((L^T x)^{ⓘ} \right)^2 > 0$$
$\Rightarrow$ A pos def.

☐

▷ Let's also show how this dets condition produces an alternating sign pattern if A is neg def.

By definition: A neg def $\Leftrightarrow$ $B := -A$ is pos def
This means $x^T B x > 0 \ \forall x \neq 0 \Rightarrow x^T(-Ax) > 0 \Rightarrow x^T A x < 0$
as we'd expect.

Here Sylvester's pos chain of principal minors looks like:

$$B = -\underbrace{\begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix}}_{A} = \underset{B_j = -A_j}{\begin{bmatrix} \overset{B_1}{-a} & -d & -e \\ -d & -b & -f \\ -e & -f & -c \end{bmatrix}} \qquad \begin{matrix} \det B_1 > 0 \\ \det B_2 > 0 \\ \det B_3 > 0 \end{matrix}$$

$0 < \det B_1 = (-1) \det A_1 \Rightarrow \det A_1 < 0$
$0 < \det B_2 = (-1)^2 \det A_2 \Rightarrow \det A_2 > 0$
$0 < \det B_3 = (-1)^3 \det A_3 \Rightarrow \det A_3 < 0$
alternating pattern

END ASIDE

9/15/2024

⑤

Now we use the sam setup stated on sheet ① and again on ③ :

$Dg_x : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ and has rank k $\Rightarrow$ $\exists$ at least one set of k cols that has non-zero det — we shall <u>assume</u> it is the <u>first set</u> $1 \dots k$ (If not, we'd have to re-index the variables)

$$Dg_x = \begin{array}{c} \begin{array}{cc} k & n-k \end{array} \\ \left[\begin{array}{c} \text{—— } Dg_1 \text{ ——} \\ \vdots \\ \text{—— } Dg_k \text{ ——} \end{array}\right] k \\ \text{call this } \boxed{B_1} \quad \boxed{B_2} \end{array}$$

So B_1^{-1} exists

$$H = D^2\mathcal{L}_{(\underline{\lambda},\underline{x})} = \begin{array}{c|c} & k \;\; H_j \;\; \overbrace{\qquad\quad}^{n} \\ \hline k & \begin{array}{c:c} 0 & -Dg_x \\ \hdashline -Dg_x^T & A \end{array} \\ n & \end{array}$$

H_j is uper left $j \times j$ submatrix.

Bordered Hessian

Robinson says $+Dg_x$ or $-Dg_x$ is irrelevant.

A is the Hessian

Robinson's Thm 3.36

- $(-1)^k \det H_j > 0 \;\; \forall j \ni 2k+1 \le j \le n+k \;\; \Rightarrow \;\; f|_M$ has local min at $\underline{x}$ (Strict?)
- $(-1)^{j-k} \det H_j > 0$ " " " $\Rightarrow$ " max "
- If neither cond holds, $\underline{x}$ is some kind of saddle pt.

It is better to first look at the examples, which follow pf.

<u>pf</u> as mentioned before, we want $A = (D^2_{x}\mathcal{L})_{\underline{\lambda},\underline{x}}$ to be pos (or neg) def on T_xM, not $\mathbb{R}^n$. Thus we must form the matrix E :

<u>Step 1</u> How can we make $v^T A v$ $v \in \mathbb{R}^n$ be restricted to $v \in T_{\underline{x}}M$? $T_{\underline{x}}M = \ker(Dg_{\underline{x}})$

$Dg_{\underline{x}} = k\left[\begin{array}{c:c} B_1 & B_2 \end{array}\right]$ (k, $n-k$) and we assumed B_1^{-1} exists. Let $v = \begin{bmatrix} x^1 \\ \vdots \\ x^k \\ x^{k+1} \\ \vdots \\ x^n \end{bmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix} a \\ b \\ c \end{matrix}}\right\} w \\ \left.\vphantom{\begin{matrix} a \\ b \\ c \end{matrix}}\right\} z \end{array}$

$Dg_{\underline{x}}(v) \overset{!}{=} 0$ Null sp can be expressed in terms of the z variables.

$$k\left[\begin{array}{c:c} \underset{k}{B_1} & \underset{n-k}{B_2} \end{array}\right] \begin{bmatrix} w \\ \hdashline z \end{bmatrix} \begin{array}{l} k \\ n-k \end{array} = 0 \;\; \Rightarrow \;\; B_1 w + B_2 z = 0$$

$$\Rightarrow \;\; w = \underbrace{\boxed{-B_1^{-1} B_2}}_{J} z \qquad k\left[B_1^{-1}\right]\left[B_2\right]k = \left[J\right]k$$

$$\Rightarrow \;\; w = Jz$$

Then
$$v = \begin{bmatrix} w \\ \hdashline z \end{bmatrix} \begin{array}{l} k \\ n-k \end{array} = \begin{array}{c} k \\ n-k \end{array} \overset{n-k}{\begin{bmatrix} J \\ \hdashline I \end{bmatrix}} \begin{bmatrix} z \end{bmatrix} n-k$$

<u>Step 2</u> Exhibit the $(n-k) \times (n-k)$ matrix E (dim T_xM is $n-k$)

$$v^T A v = \begin{bmatrix} w^T & z^T \end{bmatrix} \begin{array}{c} k \\ n-k \end{array} \overset{\begin{array}{cc} k & n-k \end{array}}{\begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{bmatrix}} \begin{bmatrix} w \\ z \end{bmatrix} = \left(\begin{bmatrix} J \\ I \end{bmatrix} z \right)^T \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{bmatrix} \begin{bmatrix} J \\ I \end{bmatrix} z$$

$$= z^T \underbrace{\boxed{\begin{bmatrix} J^T & I^T \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{bmatrix} \begin{bmatrix} J \\ I \end{bmatrix}}}_{E} z = z^T E z$$

So $v^T A v \ge 0 \Leftrightarrow z^T E z \ge 0$

But E is defined on $\mathbb{R}^{n-k}$ so we can apply Sylvester's det Thm for principal minors to show E pos/neg def →

Multiplying it out $E = J^T A_{11} J + J^T A_{12} + A_{12}^T J + A_{22}$

Step 3 Establish $\boxed{\det H = (-1)^k (\det B_1)^2 \det E}$

We must follow some fiendishly tricky obscure route to calculate principle minors in E. We will work with $H = D^2 \mathcal{L}_{(\lambda, x)} = \begin{bmatrix} 0 & B_1 & B_2 \\ B_1^T & & \\ B_2^T & & A \end{bmatrix}$

$$\det H = \det \begin{array}{c|ccc|} & k & k & n-k \\ k & 0 & B_1 & B_2 \\ k & B_1^T & A_{11} & A_{12} \\ n-k & B_2^T & A_{12}^T & A_{22} \end{array} = \underbrace{\det R}_{=1} \cdot \det H \cdot \underbrace{\det R^T}_{=1} = \det(R H R^T)$$

$$= \det \left(\underset{R}{\begin{bmatrix} I_k & 0 & 0 \\ 0 & I_k & 0 \\ 0 & J^T & I_{n-k} \end{bmatrix}} \begin{bmatrix} 0 & B_1 & B_2 \\ B_1^T & A_{11} & A_{12} \\ B_2^T & A_{12}^T & A_{22} \end{bmatrix} \begin{bmatrix} I_k & & \\ & I_k & J \\ & & I_{n-k} \end{bmatrix} \right)$$

$$= \det \left(\begin{bmatrix} I & & \\ & I & \\ & J^T & I \end{bmatrix} \begin{bmatrix} 0 & B_1 & (B_1 J + B_2) \\ B_1^T & A_{11} & (A_{11} J + A_{12}) \\ B_2^T & A_{12}^T & (A_{12}^T J + A_{22}) \end{bmatrix} \right)$$

$B_1(-B_1^{-1} B_2) + B_2 = 0$

$$= \det \begin{bmatrix} 0 & B_1 & \overbrace{(B_1 J + B)} \\ B_1^T & A_{11} & \boxed{(A_{11} J + A_{12})} \text{ call this } C \\ \underbrace{(J^T B_1^T + B_2^T)}_{0} & \underbrace{(J^T A_{11} + A_{12}^T)}_{C^T} & \underbrace{(J^T (A_{11} J + A_{12}) + A_{12}^T J + A_{22})}_{E} \end{bmatrix}$$

$$= \det \begin{bmatrix} 0 & B_1 & 0 \\ B_1^T & A_{11} & C \\ 0 & C^T & E \end{bmatrix}$$

transpose these k rows to the top

$$= (-1)^k \det \begin{bmatrix} B_1^T & A_{11} & C \\ 0 & B_1 & 0 \\ 0 & C^T & E \end{bmatrix} = (-1)^k \det(B_1^T) \det \begin{bmatrix} B_1 & 0 \\ C^T & E \end{bmatrix}$$

lower triang

$$= (-1)^k \det B_1^T \det B_1 \det E$$

$\boxed{\det H = (-1)^k (\det B_1)^2 \det E}$ Thus if we know the sign of det H, we know the sign of det E, because k fixed.

$\frac{1}{-1} = -1$ Thus $\frac{1}{(-1)^k} = (-1)^k$ $\boxed{(-1)^k \det H = \underset{\text{always pos}}{(\det B_1)^2} \det E}$

9/15/2024

⑦

Step 4 Establish $(-1)^k \det H_j = (\det B_1)^2 \det E_{j-2k} \ \forall j \quad 2k+1 \le j \le n+k$

Claim: Shaving off the last row and col of H has the effect of shaving off the last row and col of E (so we get all the upper left submats for dets)

$$H = \begin{array}{c|ccc} & k & k & n-k \\ \hline k & 0 & B_1 & B_2 \\ k & B_1^T & A_{11} & A_{12} \\ n-k & B_2^T & A_{12}^T & A_{22} \end{array}$$

Then if we trim off last row and col:

$$\begin{array}{c|ccc} & k & k & n-k-1 \\ \hline k & 0 & B_1 & \tilde{B}_2 \\ k & B_1^T & A_{11} & \tilde{A}_{12} \\ n-k-1 & \tilde{B}_2^T & \tilde{A}_{12}^T & \tilde{A}_{22} \end{array}$$

J would become $\tilde{J} = -\left[B_1^{-1}\right]^{k}_{k}\left[\tilde{B}_2\right]^{n-k-1}_{k} = \left[\tilde{J}\right]^{n-k-1}_{k}$

$$\tilde{E} = \tilde{J}^{T\,\cancel{k}}_{(n-k-1)}\left(A_{11}{}_{k}\,\tilde{J}^{n-k-1}_{\cancel{k}} + \tilde{A}_{12}{}^{n-k-1}_{\cancel{k}}\right) + \tilde{A}^{T\,\cancel{k}}_{12\,(n-k-1)}\,\tilde{J}^{n-k-1}_{\cancel{k}} + \tilde{A}_{22}{}^{n-k-1}_{(n-k-1)}$$

$\Rightarrow \tilde{E} = \left[\quad\right]^{n-k-1}_{n-k-1}$ Thus we have shaved off the last row and col of E but we must stop before E has size 0

Stop condition: $n-k-p = 1$ p = how many cols removed.

Values possible for p: $p = 0, 1, \ldots, n-k+1$

Let q be the largest col index that exists in H for a given value of p

$p = 0 \Rightarrow q = n+k$

$\vdots$

$p = n-k+1 \Rightarrow q = (n+k)-(n-k+1) = 2k+1$

so $2k+1 \le q \le n+k$

rename q as "j".

This is where we get $2k+1 \le j \le n+k$.

How do we get the index E_{j-2k}?

when j has largest value $j = n+k$ then $j-2k = (n+k)-2k = n-k$ max size for E

$\vdots$

when j has smallest value $j = 2k+1$ $j-2k = 1$ smallest size for E

$$\begin{array}{ccc} E_1, & \ldots, & E_{n-k} \\ \downarrow & & \searrow \\ j=2k+1 & & j=n+k \end{array}$$

Step 5 How do we get $(-1)^{j-k} \det H_j > 0$ as the cond for E being neg def?

we know $\det E_1$ is neg if E neg def

we have established $(-1)^k \det H_j = (\det B_1)^2 \det E_{j-2k}$

and $j = 2k+1$ makes $(-1)^k \det H_j = \underbrace{(\det B_1)^2}_{\text{pos}} \underbrace{\det E_1}_{\text{neg}}$

we must mult RHS by (-1) or $(-1)^{\text{odd power}}$ to make it pos as in thm statement

Mult both sides by $(-1)^{j-2k}$

$(-1)^{j-2k+k} \det H_j = (-1)^{j-2k} (\det B_1)^2 \det E_{j-2k}$

This coeff $(-1)^{j-2k}$ gives -1 correctly when $j = 2k+1$ and E_1 and it alternates, so this works

□

Let's work thru some of Robinson's examples

example 2 $f: \mathbb{R}^3 \longrightarrow \mathbb{R}$ $\quad g: \mathbb{R}^3 \longrightarrow \overset{k}{\mathbb{R}^1}$

$(x,y,z) \longmapsto x^2+y^2+z^2$ $\quad x,y,z \longmapsto z-xy$ we want the surf $M = g^{-1}(2)$

$n=3$

$k=1$

$\nabla f = \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix}$ $\quad \nabla g = \begin{bmatrix} -y \\ -x \\ 1 \end{bmatrix}$ we must solve:

$$2x \overset{!}{=} -\lambda y$$
$$2y = -\lambda x$$
$$2z = \lambda$$
$$z - xy = 2$$

They give us the solns: $\langle \lambda, x, y, z \rangle = \langle 4, 0, 0, 2 \rangle$

$\langle 2, 1, -1, 1 \rangle$

$\langle 2, -1, 1, 1 \rangle$

$$\mathcal{L} = f - \lambda(g-2)$$
$$= (x^2+y^2+z^2) - \lambda(z-xy) + 2\lambda$$

$$D\mathcal{L} = [\mathcal{L}_\lambda \ \mathcal{L}_x \ \mathcal{L}_y \ \mathcal{L}_z] = \left[(-z+xy+2) \ (2x+\lambda y) \ (2y+\lambda x) \ (2z-\lambda) \right]$$

$$D^2\mathcal{L} = \begin{bmatrix} \overset{H_3}{\boxed{\begin{matrix} 0 & y & x \\ y & 2 & \lambda \\ x & \lambda & 2 \end{matrix}}} & \begin{matrix} -1 \\ 0 \\ 0 \end{matrix} \\ \begin{matrix} -1 & 0 & 0 \end{matrix} & 2 \end{bmatrix} = H_4$$

$n=3 \quad k=1$ local min cond: $(-1)^{k=1} \det H_j \overset{!}{>} 0$ for all $j \ni 2k+1 \leq j \leq n+k$

$3 \leq j \leq 4$

local max cond: $(-1)^{j-k} \det H_j > 0$

$\Rightarrow H_3$ and H_4

$$\Rightarrow (-1)^{3-1} \det H_3 \overset{!}{>} 0$$
$$\boxed{\det H_3 > 0}$$

and $(-1)^{4-1} \det H_4 > 0$ or $\boxed{-\det H_4 > 0}$

$$\det H_3 = 0 - y(2y - \lambda x) + x(\lambda y - 2x)$$

For det H_4, recall formula from Strang LAAIA ch 4.3

$$\det A = \sum_{j=1}^{n} a_{ij} (-1)^{i+j} \det \underbrace{M_{ij}}_{\text{delete row } i, \text{ col } j \text{ from } A}$$ this is expanding down row i

Here we expand along row $i=4$:

$$\det H_4 = a_{41}(-1)^{4+1} \det M_{41} + 0 + 0 + 2(-1)^{4+4} \underbrace{\det M_{44}}_{\det H_3}$$
$$= (-1)(-1)^5 \det \begin{vmatrix} y & x & -1 \\ 2 & \lambda & 0 \\ \lambda & 2 & 0 \end{vmatrix} + 2 \det H_3$$
$$= (-1)(4-\lambda^2) + 2 \det H_3$$

▷ Evaluate at $\langle \lambda, x, y, z \rangle = \langle 4, 0, 0, 2 \rangle$

$$\det H_3 = \det \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 4 & 2 \end{bmatrix} = 0$$

$$\det H_4 = \det \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 2 & 4 & 0 \\ 0 & 4 & 2 & 0 \\ -1 & 0 & 0 & 2 \end{bmatrix} = -(-1) \det \begin{bmatrix} 0 & 2 & 4 \\ 0 & 4 & 2 \\ -1 & 0 & 0 \end{bmatrix} = \det \begin{vmatrix} 0 & 2 & 4 \\ 0 & 4 & 2 \\ -1 & 0 & 0 \end{vmatrix} = -1(4-16) = +12$$

But the thm wants $-\det H_3 > 0$ for local min <u>NO</u>

$-\det H_4 > 0$ <u>NO</u>

So it must be some type of Saddle pt at $\langle 4, 0, 0, 2 \rangle$

And for Local max: $\det H_3 > 0$ <u>NO</u>, $-\det H_4 > 0$ NO

Let's consider next pt ⟶

9/23/2024

Now consider $\langle \lambda, x, y, z\rangle = \langle 2, 1, -1, 1\rangle$

$$\det H_3 = +1(-2-\lambda^2) + 1(-\lambda^2-2) = -4 + -4 = -8$$ can be local min

$$\det H_4 = (-1)(4-4) + 2\det H_3 = 0 + 2(-8) = -16$$

$$-\det H_4 = +16 > 0 \Rightarrow \text{Local min}$$

Skipping doing pt $\langle 2, -1, 1, 1\rangle$

$\boxed{\text{example 3}}$

$$f(x,y,z) = z$$
$$g_1(x,y,z) = x+y+z-12$$
$$g_2(x,y,z) = x^2+y^2-z$$

$f: \mathbb{R}^3 \longrightarrow \mathbb{R}$

$g: \mathbb{R}^3 \longrightarrow \mathbb{R}^{\textcircled{2}}$

$M = g^{-1}(0)$

$\boxed{\begin{matrix} n=3 \\ k=2 \end{matrix}}$

$2k+1 \le j \le n+k$

$5 \le j \le 5$

only need H_5 $(-D^2\mathcal{L})$

$\vec{\lambda} = \begin{bmatrix} \lambda \\ \mu \end{bmatrix}$

$$\mathcal{L}(\lambda, \mu, x, y, z) = f(x,y,z) - \lambda g^{\textcircled{1}}(x,y,z) - \mu g^{\textcircled{2}}(x,y,z)$$
$$= z - \lambda(x+y+z-12) - \mu(x^2+y^2-z)$$

$$D\mathcal{L}_{(\lambda,x)} = \left[-g^{\textcircled{1}} \quad -g^{\textcircled{2}} \quad \underbrace{(f_x - \lambda g'_x - \mu g^2_x)}_{=0} \quad \underbrace{(f_y - \lambda g'_y - \mu g^2_y)}_{=0} \quad \underbrace{(f_z - \lambda g'_z - \mu g^2_z)}_{=0} \right]$$

at the critical pts (these are the Lagrange Mult eqs, but we get to that later.

$$= \left[-(x+y+z-12) \quad -(x^2+y^2-z) \quad -\lambda-2x\mu \quad -\lambda-2y\mu \quad \mu-\lambda \right]$$

$$D^2\mathcal{L}_{(\lambda,x)} = \begin{bmatrix} 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & -2x & -2y & +1 \\ -1 & -2x & -2\mu & 0 & 0 \\ -1 & -2y & 0 & -2\mu & 0 \\ -1 & 1 & 0 & 0 & 0 \end{bmatrix} = H_5$$

critical pts (Avez way):

$$Df_{\underline{x}} = \sum^{2} \lambda_i Dg^{\textcircled{i}}_{\underline{x}} \Rightarrow [f_x \; f_y \; f_z] = \lambda[g^{\textcircled{1}}_x \; g^{\textcircled{1}}_y \; g^{\textcircled{1}}_z] + \mu[g^{\textcircled{2}}_x \; g^{\textcircled{2}}_y \; g^{\textcircled{2}}_z]$$
$$[0 \; 0 \; 1] = \lambda[1 \; 1 \; 1] + \mu[2x \; 2y \; -1]$$

$\Rightarrow$ (1) $0 = \lambda + 2\mu x$

(2) $0 = \lambda + 2\mu y$

(3) $1 = \lambda - \mu \Rightarrow \lambda = 1+\mu$

subs and get

(1) $0 = 1 + \mu + 2x\mu$

(2) $-\; 0 = 1 + \mu + 2\mu y$ Subtract these eqs

$0 = 0 + 2\mu(x-y)$

so either $\mu = 0$ or $x = y$

We can rule out the $\mu = 0$ case

(1) $\Rightarrow \lambda = 0$ (3) $\Rightarrow 1 = 0 \Rightarrow\Leftarrow$

(2) $\Rightarrow \lambda = 0$

No soln's to Lag mult eqs

cont'd →

9/23/2024

⑩

Case $y = x$

$g^{①}: x + y + z = 12 \Rightarrow 2x + z = 12 \Rightarrow 2x^2 + 2x - 12 = 0$

$g^{②}: x^2 + y^2 - z = 0 \qquad 2x^2 = z$

$x^2 + x - 6 = 0$

$(x+3)(x-2) = 0$

$x = -3, 2$

if $\boxed{x=2}$ then $y = 2$

$g^{②}: z = 2x^2 = 2 \cdot 4 = 8$ $\qquad g^{①}$ yields nothing new

Lag mult 1 : $0 = 1 + \mu + 2 \cdot 2 \cdot \mu \Rightarrow 6 = 1 + 5\mu \Rightarrow \mu = {-1}/5$

Lag mult 2 : $0 = 1 + \mu + 2 \cdot 2\mu$ SAME

Lag mult 3 : $\lambda = 1 + \mu = 1 - \frac{1}{5} \Rightarrow \lambda = \frac{4}{5}$

$\boxed{\langle \lambda, \mu, x, y, z \rangle = \langle \frac{4}{5}, \frac{-1}{5}, 2, 2, 8 \rangle}$

if $\boxed{x = -3}$ $y = -3$

$z = 2(-3)^2 = 18$

$0 = 1 + \mu + 2(-3)\mu \Rightarrow \mu = 1/5$

$\lambda = 1 + \mu = 1 + \frac{1}{5} \Rightarrow \lambda = 6/5$

$\boxed{\langle \frac{6}{5}, \frac{1}{5}, -3, -3, 18 \rangle}$

Recall conds: $n = 3$, $k = 2$, only $j = 5$

LOCAL MIN: $(-1)^k \det H_j > 0$

here $(-1)^2 \det H_5 > 0$

$\Rightarrow \boxed{\det H_5 > 0}$

local MAX: $(-1)^{j-k} \det H_j > 0$

here $(-1)^{5-2} \det H_5 > 0$

$\boxed{-\det H_5 > 0}$

So all I need to do is copute $\det H_5$ and see the sign:

plug in $\langle \frac{4}{5}, \frac{-1}{5}, 2, 2, 8 \rangle$

[Compute det by Gaussian elimination]

$$\det H_5 = \det \begin{bmatrix} 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & -4 & -4 & 1 \\ -1 & -4 & 2/5 & 0 & 0 \\ -1 & -4 & 0 & 2/5 & 0 \\ -1 & 1 & 0 & 0 & 0 \end{bmatrix} \text{ transpose } = (-1)^2 \det \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ -1 & -4 & 0 & 2/5 & 0 \\ -1 & -4 & 2/5 & 0 & 0 \\ 0 & 0 & -4 & -4 & 1 \\ 0 & 0 & -1 & -1 & -1 \end{bmatrix}$$

$-4 \quad -4 \quad -4$

$R_2 \to R_1 - R_2$

$R_3 \to R_1 - R_3$

$R_5 \to R_4 - 4R_5$

$R_3 \to R_2 - R_3$

[redacted]

$$= \det \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & +5 & 0 & -2/5 & 0 \\ 0 & +5 & -2/5 & 0 & 0 \\ 0 & 0 & -4 & -4 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} \leftarrow -\frac{2}{5} \; -\frac{2}{5} \; \frac{1}{10} \quad = \det \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & -2/5 & 0 \\ 0 & 0 & 2/5 & -2/5 & 0 \\ 0 & 0 & 0 & -4/5 & \frac{1}{10} \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

$R_4 \to R_3 + \frac{2/5}{4} R_4$

$\det = (-1)(\cancel{5})(\frac{2}{\cancel{5}})(-\frac{4}{\cancel{5}})(\cancel{5}) = +8$

local MIN

Robinson gets $\det H_5 = 20$ so local MIN

There is an arithmetic mistake somewhere, but I am not looking further for it.

▷ I am not doing the calculation for $\langle 6/5, 1/5, -3, -3, 18 \rangle$

It is supposed to be local Max.

☐