SECOND-ORDER CONDITIONS AND QUADRATIC FORMS WHEN THERE ARE CONSTRAINTS

RECALL THAT AT A POINT $\bar{x} \in \mathbb{R}^n$ WE APPROXIMATE THE CHANGE IN THE VALUE OF A FUNCTION $f : \mathbb{R}^n \to \mathbb{R}$ THAT RESULTS FROM A CHANGE ("DISPLACEMENT") $\Delta x \in \mathbb{R}^n$ BY A TAYLOR POLYNOMIAL. USING JUST THE FIRST-DEGREE (LINEAR) TAYLOR POLYNOMIAL, WE HAVE

$$\Delta f \approx \nabla f(\bar{x}) \cdot \Delta x = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\bar{x})\, \Delta x_i ,$$

AND THIS MUST BE ZERO FOR ALL $\Delta x \in \mathbb{R}^n$ IF $\bar{x}$ MAXIMIZES OR MINIMIZES f, i.e., $\nabla f(\bar{x}) = \underline{0}$. $\bar{x}$ IS A CRITICAL POINT

WE USE THE SECOND-DEGREE TAYLOR POLYNOMIAL AS OUR APPROXIMATION IF WE WANT TO DETERMINE WHETHER A CRITICAL POINT $\bar{x}$ IS A LOCAL MAXIMUM OR MINIMUM:

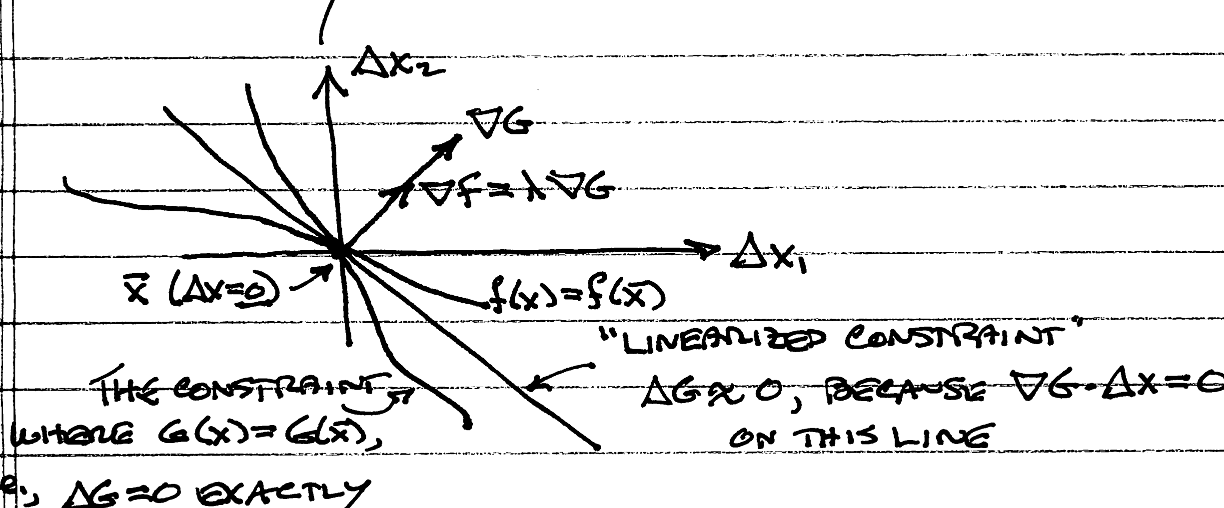
$$\begin{aligned} \Delta f &\approx \nabla f \cdot \Delta x + \tfrac{1}{2} \Delta x D^2 f \Delta x \\ &= \tfrac{1}{2} \Delta x Df^2 \Delta x, \text{ SINCE } \nabla f = \underline{0} \text{ AT } \bar{x}. \end{aligned}$$

FOR $\bar{x}$ TO BE A LOCAL MAXIMUM THE QUADRATIC FORM $\Delta x^T D^2 f \Delta x$ HAS TO BE NEGATIVE DEFINITE (THE SUFFICIENT CONDITION, WITH $\nabla f = \underline{0}$) OR NEGATIVE SEMIDEFINITE (THE NECESSARY CONDITION), AND SIMILARLY FOR POSITIVE DEFINITE/SEMIDEFINITE IF $\bar{x}$ IS TO BE A LOCAL MINIMUM OF f.

In order that $\bar{x}$ be a maximum or minimum of f <u>subject to a constraint</u> $G(x)=0$, we need to have $\nabla f - \lambda \nabla G = \underline{0}$ (i.e., $\nabla f = \lambda \nabla G$) at $\bar{x}$ for some λ, and the quadratic form $\Delta x\, D f^2 \Delta x$ must be positive or negative definite/semidefinite <u>subject to the constraint</u> $G(x)=G(\bar{x})$ – i.e., subject to the linear approximation

$$\Delta G \approx \nabla G \cdot \Delta x = 0 = \sum_{i=1}^{n} \frac{\partial G}{\partial x_i} \Delta x_i = 0.$$

The quadratic form $\Delta x\, D f^2 \Delta x$ doesn't need to be zero for <u>all</u> $\Delta x \neq \underline{0} \in \mathbb{R}^n$; it only needs to be zero for those Δx that satisfy the linear constraint $\nabla G \cdot \Delta x = 0$.



So we want to determine conditions under which a quadratic form will be positive or negative definite subject to a linear constraint — or to several linear constraints, if we want to know about optimization subject to multiple constraints.

QUADRATIC FORMS

SUBJECT TO LINEAR CONSTRAINTS

A CENTRAL TOOL FOR WORKING WITH SECOND-ORDER (CURVATURE) CONDITIONS WHEN THERE ARE CONSTRAINTS IS <u>BORDERED MATRICES</u> AND THEIR DETERMINANTS. LET'S START OFF BY DEFINING BORDERED MATRICES AND (FOR THE 2x2 CASE WITH A SINGLE CONSTRAINT) EVALUATING THEIR DETERMINANTS.

<u>EXAMPLE:</u>

LET A BE A 2x2 MATRIX $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ AND LET $b = (b_1, b_2) \in \mathbb{R}^2$. [IN GENERAL, WE WILL HAVE AN $n \times n$ MATRIX A AND m n-TUPLES (OR $n \times 1$ MATRICES) $b^i \in \mathbb{R}^n$.] ($m<n$) THE ASSOCIATED <u>BORDERED MATRIX</u> IS

$$\mathbb{B} = \begin{bmatrix} 0 & b_1 & b_2 \\ b_1 & a_{11} & a_{12} \\ b_2 & a_{21} & a_{22} \end{bmatrix}$$ WHICH WE COULD WRITE AS $$\left[\begin{array}{c|c} 0 & b \\ \hline b & A \end{array}\right].$$

ITS DETERMINANT IS

$$|\mathbb{B}| = \begin{vmatrix} 0 & b_1 & b_2 \\ b_1 & a_{11} & a_{12} \\ b_2 & a_{21} & a_{22} \end{vmatrix} = a_{12}b_1b_2 + a_{21}b_1b_2 - a_{11}b_2b_2 - a_{22}b_1b_1.$$

WE USE THESE CONCEPTS MOSTLY WITH SYMMETRIC MATRICES A, WHERE WE WOULD HAVE

$$|\mathbb{B}| = 2a_{12}b_1b_2 - a_{11}b_2^2 - a_{22}b_1^2.$$

NOTE THAT $|\mathbb{B}|$ IS THE SAME IF WE PUT THE BORDER AT THE RIGHT AND BOTTOM: *

$$|\mathbb{B}| = \left|\begin{array}{cc|c} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ \hline b_1 & b_2 & 0 \end{array}\right| .$$

IN THE GENERAL m-CONSTRAINT, n-VARIABLE CASE, WE HAVE

$$\mathbb{B} = \left[\begin{array}{c|c} \underset{m\times m}{0} & \underset{m\times n}{B} \\ \hline \underset{n\times m}{B} & \underset{n\times n}{A} \end{array}\right] ; \quad \mathbb{B} \text{ IS } (m+n)\times(m+n).$$

WE'RE GOING TO FOCUS ON THE $m=1$ (ONE CONSTRAINT) CASE.

* BOTH WAYS OF DOING IT ARE COMMON — BORDER ON THE LEFT AND TOP, AND BORDER ON THE RIGHT AND BOTTOM. IN SOME 1ST-YEAR TEXTBOOKS:

LEFT & TOP: JEHLE & RENY; SIMON & BLUME; DE LA FUENTE.

RIGHT & BOTTOM: MAS COLELL, WHINSTON, AND GREEN; VARIAN.

EXAMPLE 16.6 IN S&B:

THIS IS AN EXAMPLE OF A QUADRATIC FORM THAT'S INDEFINITE ON THE ENTIRE SPACE ($\mathbb{R}^2$ IN THIS CASE), BUT IS POSITIVE DEFINITE OR NEGATIVE DEFINITE ON EVERY LINE THROUGH THE ORIGIN IN $\mathbb{R}^2$ EXCEPT TWO (ON EACH OF WHICH IT'S IDENTICALLY ZERO).

$$Q(x_1,x_2) = x_1^2 - x_2^2 = x^T A x \quad \text{FOR } A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

A IS INDEFINITE: $\exists x \in \mathbb{R}^2 : x^T A x > 0$

AND $\exists x \in \mathbb{R}^2 : x^T A x < 0$.

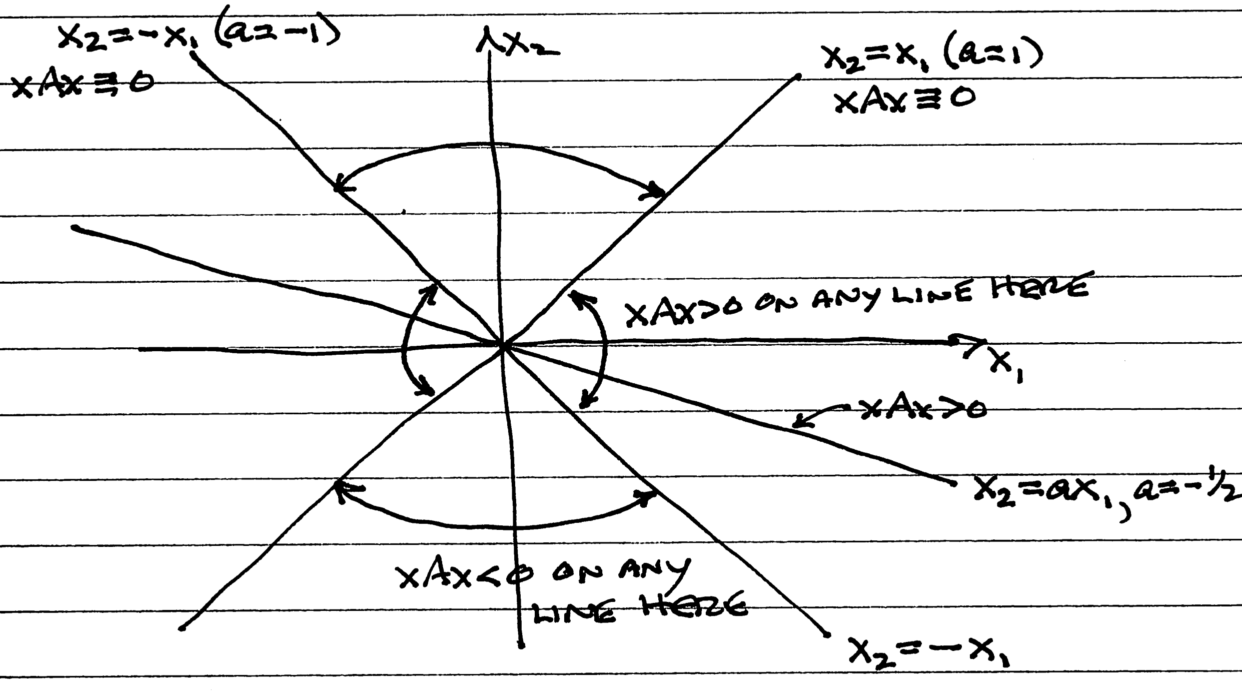
IN PARTICULAR: A LINE

IF $x_2 = a x_1$, THEN $x^T A x = x_1^2 - a^2 x_1^2 = (1-a^2) x_1^2$

SO $x^T A x > 0$, $\forall x$ s.t. $x_2 = a x_1$ IF $|a| < 1$ (EXCEPT $x = (0,0)$)

AND $x^T A x < 0$, $\forall x$ s.t. $x_2 = a x_1$ IF $|a| > 1$ (EXCEPT $x = (0,0)$).

AND $x^T A x = 0$, $\forall x$ s.t. $x_2 = a x_1$ IF $|a| = 1$.



LET'S SEE WHAT HAPPENS FOR AN ARBITRARY QUADRATIC FORM IN $\mathbb{R}^2$, $Q(x_1, x_2) = x^T A x$, IF WE RESTRICT OURSELVES TO A LINE

$$b_1 x_1 + b_2 x_2 = 0, \quad \text{i.e., } b \cdot x = 0.$$

AT LEAST ONE OF THE COEFFICIENTS b_1, b_2 MUST BE NON-ZERO; WLOG LET'S SAY $b_2 \neq 0$. THEN WE CAN WRITE

$$x_2 = -\frac{b_1}{b_2} x_1$$

AND SUBSTITUTE THIS INTO THE QUADRATIC FORM:

$$Q(x_1, x_2) = x^T A x = a_{11} x_1^2 + a_{12} x_1 x_2 + a_{21} x_2 x_1 + a_{22} x_2^2$$

$$= a_{11} x_1^2 + 2a_{12} x_1 x_2 + a_{22} x_2^2, \text{ BECAUSE } A \text{ IS SYMMETRIC}$$

$$= a_{11} x_1^2 + 2a_{12} x_1 \left(-\frac{b_1}{b_2} x_1\right) + a_{22} \left(-\frac{b_1}{b_2} x_1\right)^2$$

$$= a_{11} x_1^2 - 2\frac{b_1}{b_2} a_{12} x_1^2 + \frac{b_1^2}{b_2^2} a_{22} x_1^2$$

$$= \frac{1}{b_2^2} x_1^2 \left[a_{11} b_2^2 + a_{22} b_1^2 - 2a_{12} b_1 b_2\right].$$

$$= \frac{1}{b_2^2} x_1^2 (-1) |\mathbb{B}|, \text{ WHERE } \mathbb{B} \text{ IS THE BORDERED MATRIX } \left[\begin{array}{c|c} 0 & b \\ \hline b & A \end{array}\right].$$

IF $b_2 = 0$ WE MUST HAVE $b_1 \neq 0$, IN WHICH CASE

$$x^T A x = \frac{1}{b_1^2} x_2^2 (-1) |\mathbb{B}|.$$

FOR A 2x2 MATRIX A AND ONE LINEAR CONSTRAINT, WE'VE PROVED THE FOLLOWING THEOREM:

THEOREM: ON THE LINE $b_1x_1 + b_2x_2 = 0$ IN $\mathbb{R}^2$ THE QUADRATIC FORM $x^T A x$ IS

POSITIVE DEFINITE IF AND ONLY IF $|\mathbb{B}| < 0$,

NEGATIVE DEFINITE IF AND ONLY IF $|\mathbb{B}| > 0$.

IN ORDER TO GENERALIZE THIS RESULT TO n VARIABLES AND ONE CONSTRAINT, OR TO n VARIABLES AND m CONSTRAINTS $(m < n)$, WE NEED TO WORK WITH THE PRINCIPAL MINORS OF THE MATRIX $\mathbb{B}$, AS WE DID WITH THE MINORS OF A WHEN WE STUDIED QUADRATIC FORMS THAT WEREN'T CONSTRAINED.

AS BEFORE, WE'LL USE THE NOTATION $\mathbb{B}_k$ FOR PRINCIPAL SUBMATRICES OF ORDER k (i.e., WITH k ROWS AND COLUMNS). THEREFORE THE LARGEST PRINCIPAL SUBMATRIX OF $\mathbb{B}$ — i.e., $\mathbb{B}$ ITSELF — IS $\mathbb{B}_{m+n}$, BECAUSE OF THE m $m \times n$ BORDERS. PRINCIPAL MINORS OF $\mathbb{B}$ ARE THE DETERMINANTS OF THE PRINCIPAL SUBMATRICES.

AS IN THE UNCONSTRAINED CASE, WE'RE GOING TO BE USING THE <u>LEADING</u> PRINCIPAL MINORS OF $\mathbb{B}$. I'VE FOUND THAT FOR BORDERED MATRICES IT'S HELPFUL TO HAVE ANOTHER, ALTERNATIVE NOTATION FOR THE LEADING PRINCIPAL MINORS:

NOTATION: FOR A BORDERED MATRIX

$$\mathbb{B} = \left[\begin{array}{c|c} 0 & B \\ \hline B^T & A \end{array}\right],$$ WHERE A IS $n \times n$ AND B IS $m \times n$,

LET $\mathbb{B}^r$ DENOTE THE PRINCIPAL SUBMATRIX IN WHICH THE LEADING (i.e., "NORTHWEST") r ROWS AND COLUMNS OF A ARE RETAINED, ← r FOR "RETAINED" ALONG WITH THE FIRST r COLUMNS OF B, THE FIRST r ROWS OF B^T, AND THE ENTIRE $m \times m$ MATRIX DENOTED BY 0 ABOVE. THUS, ONLY THE $n-r$ RIGHTMOST COLUMNS AND THE $n-r$ BOTTOM ROWS OF $\mathbb{B}$ ARE DELETED. NOTE THAT $\mathbb{B}^r$ HAS $m+r$ ROWS AND COLUMNS, SO IT'S THE LEADING PRINCIPAL SUBMATRIX OF $\mathbb{B}$ OF ORDER $k = m+r$. (AS ALWAYS, A MINOR IS THE DETERMINANT OF A SUBMATRIX.)

THM 16.4 IN S&B

THEOREM: LET A BE AN $n \times n$ SYMMETRIC MATRIX ← WHERE $m<n$ AND rank $B=m$ AND LET B BE AN $m \times n$ MATRIX. ON THE SUBSPACE $C = \{x \in \mathbb{R}^n \mid Bx = \underline{0}\}$, THE QUADRATIC FORM $x^T A x$ IS

(a) POSITIVE DEFINITE IF $(-1)^m |\mathbb{B}^r| > 0$ FOR $r = m+1, \dots, n$;

(b) NEGATIVE DEFINITE IF $(-1)^r |\mathbb{B}^r| > 0$ FOR $r = m+1, \dots, n$.

← SHOULD BE IF AND ONLY IF

SO WE CONSIDER ONLY THE LEADING PRINCIPAL MINORS OF $\mathbb{B}$ OF ORDER $k = 2m+1, \dots, \cancel{n}$ — i.e., THOSE OF ORDER LARGER THAN $2m$.
↑ $m+n$

FOR EXAMPLE:

m CONSTRAINTS, n VARIABLES → m,n; ROWS AND COLUMNS OF $\mathbb{B}$ → $m+n$; THE MINORS WE CONSIDER → r, k

m,n	$m+n$	r	k
1,2	3	2	3
1,3	4	2,3	3,4
1,4	5	2,3,4	3,4,5
2,3	5	3	5
2,4	6	3,4	5,6

FOR EXAMPLE, IF $m=1$ AND $n=4$:

$$\left.\begin{array}{l} r=m+1=2=n-2 \\ k=2m+1=3 \end{array}\right\} \quad |\mathbb{B}^2| = \begin{vmatrix} 0 & b_1 & b_2 \\ b_1 & a_{11} & a_{12} \\ b_2 & a_{21} & a_{22} \end{vmatrix} \begin{array}{l} < 0 \text{ POS. DEF.} \\ > 0 \text{ NEG. DEF.} \end{array} \qquad (-1)^{\overset{r}{2}} = 1$$

$$\left.\begin{array}{l} r=m+2=3=n-1 \\ k=2m+2=4 \end{array}\right\} \quad |\mathbb{B}^3| = \begin{vmatrix} 0 & b_1 & b_2 & b_3 \\ b_1 & a_{11} & a_{12} & a_{13} \\ b_2 & a_{21} & a_{22} & a_{23} \\ b_3 & a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{array}{l} < 0 \text{ POS. DEF.} \\ < 0 \text{ NEG. DEF.} \end{array} \qquad (-1)^{\overset{r}{3}} = -1$$

$$\left.\begin{array}{l} r=m+3=4=n \\ k=2m+3=5 \\ \quad =m+n \end{array}\right\} \quad |\mathbb{B}^4| = \left|\begin{array}{c|c} 0 & B \\ \hline B^T & \underset{\text{A is } 4\times 4}{A} \end{array}\right| \begin{array}{l} < 0 \text{ POS. DEF.} \\ > 0 \text{ NEG. DEF.} \end{array} \qquad (-1)^{\overset{r}{4}} = 1$$

EXAMPLE: $m=2,\ n=4$

$r=m+1=3=n-1$
$k=2m+1=5$

$r=3$: $\quad \mathbb{B}^3 = \left|\begin{array}{cc:ccc} 0 & 0 & b_{11} & b_{12} & b_{13} \\ 0 & 0 & b_{21} & b_{22} & b_{23} \\ \hdashline b_{11} & b_{21} & a_{11} & a_{12} & a_{13} \\ b_{12} & b_{22} & a_{12} & a_{22} & a_{23} \\ b_{13} & b_{23} & a_{13} & a_{23} & a_{33} \end{array}\right|$

>0 FOR POS. DEF.
$(-1)^2|\mathbb{B}^3|>0$

<0 FOR NEG. DEF.
$(-1)^3|\mathbb{B}^3|<0$

$r=m+2=4=n$
$k=2m+2=6$
$=m+n$

$r=4$: $\quad \mathbb{B}^4 = \left|\begin{array}{cc:cccc} 0 & 0 & b_{11} & b_{12} & b_{13} & b_{14} \\ 0 & 0 & b_{21} & b_{22} & b_{23} & b_{24} \\ \hdashline b_{11} & b_{21} & & & & \\ b_{12} & b_{22} & & A & & \\ b_{13} & b_{23} & & & & \\ b_{14} & b_{24} & & & & \end{array}\right|$

>0 FOR POS. DEF.
$(-1)^2|\mathbb{B}^4|>0$

>0 FOR NEG. DEF.
$(-1)^4|\mathbb{B}^4|>0$

THE CONDITIONS FOR A TO BE POS. DEF. / NEG. DEF. SUBJECT TO $Bx=\underline{0}$ INCLUDE ~~Fo~~ JUST THESE TWO DETERMINANTS: $r=m+1,\dots,n$

i.e., $r=3,4$.

SEMIDEFINITENESS:

THE FOLLOWING THEOREM GIVES CONDITIONS FOR xAx TO BE POSITIVE OR NEGATIVE SEMIDEFINITE SUBJECT TO CONSTRAINT. THE RELATION OF THESE CONDITIONS TO THE CONDITIONS WE'VE JUST DEVELOPED FOR xAx TO BE POSITIVE OR NEGATIVE DEFINITE SUBJECT TO CONSTRAINT IS PARALLEL TO THE RELATION BETWEEN THE DEFINITESS AND SEMIDEFINITENESS CONDITIONS IN THE UNCONSTRAINED CASE.

THEOREM: THE QUADRATIC FORM xAx IS POSITIVE OR NEGATIVE SEMIDEFINITE ON THE SET $\{x \in \mathbb{R}^n \mid Bx = \underline{0}\}$ IF AND ONLY IF THE CONDITIONS (a) AND (b) ARE IN THE PRECEDING THEOREM CHANGED AS FOLLOWS:

(i) ALL INEQUALITIES ARE CHANGED TO WEAK INEQUALITIES;

(ii) THE CONDITIONS ARE EXPANDED TO INCLUDE ALL BORDER-PRESERVING PRINCIPAL MINORS FOR $r = m+1, \ldots, n$ — i.e., TO INCLUDE ALL PERMUTATIONS OF THE ROWS AND COLUMNS OF A.

IN OTHER WORDS, JUST AS IN THE UNCONSTRAINED CASE, WE CAN ENSURE DEFINITENESS (i.e., WE HAVE SUFFICIENCY) WITH INEQUALITIES INVOLVING JUST THE LEADING PRINCIPAL MINORS (THE BORDER-PRESERVING ONES, HERE), BUT FOR SEMIDEFINITENESS ALL THE APPROPRIATE-ORDER PRINCIPAL MINORS MUST ~~HA~~ SATISFY THE INEQUALITIES. ALL MINORS ARE NEEDED BECAUSE THE INEQUALITIES ARE CHANGED FROM STRICT TO WEAK.

Optimization with Constraints

We can now add second-order conditions to our first-order conditions for a local maximum or minimum of a function, subject to constraints that are equations.

Theorem: Assume that $f : \mathbb{R}^n \to \mathbb{R}$ and $G^i : \mathbb{R}^n \to \mathbb{R}$ $(i = 1, \ldots, m)$ are C^2-functions; let $\overline{\mathbf{x}} \in \mathbb{R}^n$ be a point at which $\nabla f(\overline{\mathbf{x}}) \neq \mathbf{0}$, and assume that

$$G^i(\overline{\mathbf{x}}) = c_i \ (i = 1, \ldots, m) \ \text{ and } \ \nabla f(\overline{\mathbf{x}}) = \lambda_1 \nabla G^1(\overline{\mathbf{x}}) + \cdots + \lambda_m \nabla G^m(\overline{\mathbf{x}})$$

for some nonzero *Lagrange multipliers* $\lambda_1, \ldots, \lambda_m$ — *i.e.*, the first-order conditions for a constrained extremum of f are satisfied at $\overline{\mathbf{x}}$. Then the second-order conditions necessary or sufficient for $\overline{\mathbf{x}} \in \mathbb{R}^n$ to be a local maximum or a local minimum point of f, subject to the constraints $G^i(\mathbf{x}) = c_i$ $(i = 1, \ldots, m)$, are the corresponding conditions for the quadratic form $\Delta\mathbf{x} H(\overline{\mathbf{x}}) \Delta\mathbf{x}$ to be negative or positive definite or semidefinite subject to the homogeneous linear constraints $\nabla G^i(\overline{\mathbf{x}}) \Delta\mathbf{x} = 0$ $(i = 1, \ldots, m)$:

For $\overline{\mathbf{x}}$ to be a local maximum point of f subject to the constraints:
 Sufficient: $\Delta\mathbf{x} H(\overline{\mathbf{x}}) \Delta\mathbf{x}$ is negative definite subject to $\nabla G^i \Delta\mathbf{x} = 0$ $(i = 1, \ldots, m)$
 Necessary: $\Delta\mathbf{x} H(\overline{\mathbf{x}}) \Delta\mathbf{x}$ is negative semidefinite subject to $\nabla G^i \Delta\mathbf{x} = 0$ $(i = 1, \ldots, m)$

For $\overline{\mathbf{x}}$ to be a local minimum point of f subject to the constraints:
 Sufficient: $\Delta\mathbf{x} H(\overline{\mathbf{x}}) \Delta\mathbf{x}$ is positive definite subject to $\nabla G^i \Delta\mathbf{x} = 0$ $(i = 1, \ldots, m)$
 Necessary: $\Delta\mathbf{x} H(\overline{\mathbf{x}}) \Delta\mathbf{x}$ is positive semidefinite subject to $\nabla G^i \Delta\mathbf{x} = 0$ $(i = 1, \ldots, m)$,

where $H(\overline{\mathbf{x}})$ is the Hessian matrix $D^2 f$ evaluated at $\overline{\mathbf{x}}$ and where each ∇G^i is evaluated at $\overline{\mathbf{x}}$.

In other words, in the bordered symmetric matrices $\mathbb{B}^r$ in the Quadratic Forms Theorems

$$\text{we replace each } a_{ij} \text{ with } \frac{\partial^2 f}{\partial x_i \partial x_j}(\overline{\mathbf{x}}) \quad \text{and each } \ b_{ij} \text{ with } \frac{\partial G^i}{\partial x_j}(\overline{\mathbf{x}}),$$

and then the conditions on the determinants $\left|\mathbb{B}^r\right|$ for the quadratic form to be negative or positive definite or semidefinite subject to constraints become the conditions for $\overline{\mathbf{x}}$ to be a maximum or minimum point of f subject to constraints.