APM462
Lecture Notes

Yuchen Wang

December 28, 2019

Contents

1 Matrix Calculus

Row v.s. Column Vector Our default rule is that every vector is a column vector unless explicitly stated otherwise.
This is also known as the numerator layout.
Special case: For $f : \mathbb{R}^n \to \mathbb{R}$, Df is a $1 \times n$ matrix or row vector.

1.1 Matrix Multiplication

Definition 1.1.1 Let A be $m \times n$, and B be $n \times p$, and let the product AB be

$$C = AB$$

then C is a $m \times p$ matrix, with element (i, j) given by

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

for all $i = 1, 2, \ldots, m, j = 1, 2, \ldots, p$.

Proposition 1.1.2 Let A be $m \times n$, and x be $n \times 1$, then the typical element of the product

$$z = Ax$$

is given by

$$z_i = \sum_{k=1}^{n} a_{ik} x_k$$

for all $i = 1, 2, \ldots, m$.
Similarly, let y be $m \times 1$, then the typical element of the product

$$z^T = y^T A$$

is given by

$$z_i^T = \sum_{k=1}^{n} a_{ki} y_k$$

for all $i = 1, 2, \ldots, n$.
Finally, the scalar resulting from the product

$$\alpha = y^T A x$$

is given by

$$\alpha = \sum_{j=1}^{m} \sum_{k=1}^{n} a_{jk} y_i x_k$$

1.2 Partitioned Matrices

Proposition 1.2.1 Let A be a square, nonsingular matrix of order m. Partition A as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

so that A_{11} and A_{22} are invertible.
Then

$$A^{-1} = \begin{bmatrix} (A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1} & -A_{11}^{-1} A_{12} (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} \\ -A_{22}^{-1} A_{21} (A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1} & (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} \end{bmatrix}$$

proof:
Direct multiplication of the proposed A^{-1} and A yields

$$A^{-1} A = I$$

■

1.3 Matrix Differentiation

Proposition 1.3.1

$$\frac{\partial A}{\partial x} = \frac{\partial A^T}{\partial x}$$

Proposition 1.3.2 Let

$$y = Ax$$

where y is $m \times 1$, x is $n \times 1$, A is $m \times n$, and A does not depend on x. Suppose that x is a function of the vector z, while A is independent of z. Then

$$\frac{\partial y}{\partial z} = A\frac{\partial x}{\partial z}$$

Proposition 1.3.3 Let the scalar α be defined by

$$\alpha = y^T Ax$$

where y is $m \times 1$, x is $n \times 1$, A is $m \times n$, and A is independent of x and y, then

$$\frac{\partial \alpha}{\partial x} = y^T A$$

and

$$\frac{\partial \alpha}{\partial y} = x^T A^T$$

Proposition 1.3.4 For the special case where the scalar α is given by the quadratic form

$$\alpha = x^T Ax$$

where x is $n \times 1$, A is $n \times n$, and A does not depend on x, then

$$\frac{\partial \alpha}{\partial x} = x^T(A + A^T)$$

proof:
By definition

$$\alpha = \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}x_i x_j$$

Differentiating with respect to the kth element of x we have

$$\frac{\partial \alpha}{\partial x_k} = \sum_{j=1}^{n} a_{kj}x_J + \sum_{i=1}^{n} a_{ik}x_i$$

for all $k = 1, 2, \ldots, n$, and consequently,

$$\frac{\partial \alpha}{\partial x} = x^T A^T + x^T A = x^T(A^T + A)$$

∎

Proposition 1.3.4 For the special case where A is a symmetric matrix and

$$\alpha = x^T Ax$$

where x is $n \times 1$, A is $n \times n$, and A does not depend on x, then

$$\frac{\partial \alpha}{\partial x} = 2x^T A$$

Proposition 1.3.5 Let the scalar α be defined by

$$\alpha = y^T x$$

where y is $n \times 1$, x is $n \times 1$, and both y and x are functions of the vector z. Then

$$\frac{\partial \alpha}{\partial z} = x^T \frac{\partial y}{\partial z} + y^T \frac{\partial x}{\partial z}$$

Proposition 1.3.6 Let the scalar α be defined by

$$\alpha = x^T x$$

where x is $n \times 1$, and x is a functions of the vector z. Then

$$\frac{\partial \alpha}{\partial z} = 2x^T \frac{\partial y}{\partial z}$$

Proposition 1.3.7 Let the scalar α be defined by

$$\alpha = y^T A x$$

where y is $m \times 1$, A is $m \times n$, x is $n \times 1$, and both y and x are functions of the vector z, while A does not depend on z. Then

$$\frac{\partial \alpha}{\partial z} = x^T A^T \frac{\partial y}{\partial z} + y^T A \frac{\partial x}{\partial z}$$

Proposition 1.3.8 Let A be an invertible, $m \times m$ matrix whose elements are functions of the scalar parameter α. Then

$$\frac{\partial A^{-1}}{\partial \alpha} = -A^{-1} \frac{\partial A}{\partial \alpha} A^{-1}$$

proof:
Start with the definition of the inverse

$$A^{-1} A = I$$

and differentiate, yielding

$$A^{-1} \frac{\partial A}{\partial \alpha} + \frac{\partial A^{-1}}{\partial \alpha} A = 0$$

rearranging the terms yields

$$\frac{\partial A^{-1}}{\partial \alpha} = -A^{-1} \frac{\partial A}{\partial \alpha} A^{-1}$$

■

Vector-by-vector Differentiation Identities 1.3.9

Young's Theorem 1.3.10 i.e. Symmetry of second derivatives

$$[\nabla_{xy} f(x, y)]^T = \nabla_{yx} f(x, y)$$

proof:
This is straightforward by writing out the elements of the matrix. ■

2 Second-year Calculus Review

functions $\mathbb{R} \to \mathbb{R}$

Condition	Expression	Numerator layout, i.e. by $\mathbf{y}$ and $\mathbf{x}^\mathsf{T}$	Denominator layout, i.e. by $\mathbf{y}^\mathsf{T}$ and $\mathbf{x}$
$\mathbf{a}$ is not a function of $\mathbf{x}$	$\frac{\partial \mathbf{a}}{\partial \mathbf{x}} =$	$\mathbf{0}$	
	$\frac{\partial \mathbf{x}}{\partial \mathbf{x}} =$	$\mathbf{I}$	
$\mathbf{A}$ is not a function of $\mathbf{x}$	$\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} =$	$\mathbf{A}$	$\mathbf{A}^\top$
$\mathbf{A}$ is not a function of $\mathbf{x}$	$\frac{\partial \mathbf{x}^\top \mathbf{A}}{\partial \mathbf{x}} =$	$\mathbf{A}^\top$	$\mathbf{A}$
a is not a function of $\mathbf{x}$, $\mathbf{u} = \mathbf{u}(\mathbf{x})$	$\frac{\partial a\mathbf{u}}{\partial \mathbf{x}} =$	$a\frac{\partial \mathbf{u}}{\partial \mathbf{x}}$	
$v = v(\mathbf{x})$, $\mathbf{u} = \mathbf{u}(\mathbf{x})$	$\frac{\partial v\mathbf{u}}{\partial \mathbf{x}} =$	$v\frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \mathbf{u}\frac{\partial v}{\partial \mathbf{x}}$	$v\frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \frac{\partial v}{\partial \mathbf{x}}\mathbf{u}^\top$
$\mathbf{A}$ is not a function of $\mathbf{x}$, $\mathbf{u} = \mathbf{u}(\mathbf{x})$	$\frac{\partial \mathbf{A}\mathbf{u}}{\partial \mathbf{x}} =$	$\mathbf{A}\frac{\partial \mathbf{u}}{\partial \mathbf{x}}$	$\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\mathbf{A}^\top$
$\mathbf{u} = \mathbf{u}(\mathbf{x})$, $\mathbf{v} = \mathbf{v}(\mathbf{x})$	$\frac{\partial (\mathbf{u}+\mathbf{v})}{\partial \mathbf{x}} =$	$\frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$	
$\mathbf{u} = \mathbf{u}(\mathbf{x})$	$\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{x}} =$	$\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial \mathbf{x}}$	$\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}$
$\mathbf{u} = \mathbf{u}(\mathbf{x})$	$\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{u}))}{\partial \mathbf{x}} =$	$\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}}\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial \mathbf{x}}$	$\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}}$

2.1 Mean Value Theorem in 1 Dimension

$g \in C^1$ on $\mathbb{R}$

$$\frac{g(x+h)-g(x)}{h} = g'(x+\theta h)$$

where $\theta \in (0,1)$
Or equivalently,

$$g(x+h) = g(x) + hg'(x+\theta h)$$

2.2 1st Order Taylor Approximation

$g \in C^1$ on $\mathbb{R}$

$$g(x+h) = g(x) + hg'(x) + o(h)$$

where $o(h)$ is "little o" of h, the error term.
Say a function $f(h) = o(h)$, this means $\lim_{h\to 0} \frac{f(h)}{h} = 0$
For example, for $f(h) = h^2$, we can say $f(h) = o(h)$,
since $\lim_{h\to 0} \frac{f(h)}{h} = \lim_{h\to 0} \frac{h^2}{h} = \lim_{h\to 0} h = 0$
proof: (Use MVT):
WTS : $g(x+h) - g(x) - hg'(x) = o(h)$

$$\begin{aligned}
\lim_{h\to 0} \frac{[g(x+h)-g(x)] - hg'(x)}{h} &= \lim_{h\to 0} \frac{[hg'(x+\theta h)] - hg'(x)}{h} \\
&= \lim_{h\to 0} g'(x+\theta h) - g'(x) \\
&= \lim_{h\to 0} g'(x) - g'(x) \\
&= 0
\end{aligned}$$

■

2.3 2nd Order Mean Value Theorem

$g \in C^2$ on $\mathbb{R}$

$$g(x+h) = g(x) + hg'(x) + \frac{h^2}{2}g'(x+\theta h)$$

for some $\theta \in (0,1)$

proof:
WTS: $g(x+h) - g(x) - hg'(x) - \frac{h^2}{2}g''(x) = o(h^2)$

$$\begin{aligned}
\lim_{h\to 0}\frac{g(x+h) - g(x) - hg'(x) - \frac{h^2}{2}g''(x)}{h^2} &= \lim_{h\to 0}\frac{[\frac{h^2}{2}g'(x+\theta h)] - \frac{h^2}{2}g''(x)}{h^2} \\
&= \lim_{h\to 0}\frac{1}{2}(g''(x+\theta h) - g''(x)) \\
&= \lim_{h\to 0}\frac{1}{2}(g''(x) - g''(x)) \\
&= 0
\end{aligned}$$

■

multivariate functions: $\mathbb{R}^n \to \mathbb{R}$

2.4 Recall: Definition of gradient

Gradient of $f : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$ (denoted $\nabla f(x)$) if exists is a vector characterized by the property:

$$\lim_{\mathbf{v}\to\mathbf{0}}\frac{f(\mathbf{x}+\mathbf{v}) - f(\mathbf{x}) - \nabla f(\mathbf{x})\cdot\mathbf{v}}{||\mathbf{v}||} = 0$$

In Cartesian coordinates, $\nabla f(\mathbf{x}) = (\frac{\partial f}{\partial \mathbf{x}_1}(\mathbf{x}), \ldots, \frac{\partial f}{\partial \mathbf{x}_n}(\mathbf{x}))$

2.5 Mean Value Theorem in n dimension

$f \in C^1$ on $\mathbb{R}^n$, then for any $\mathbf{x}, \mathbf{v} \in \mathbb{R}^n$,

$$f(\mathbf{x}+\mathbf{v}) = f(\mathbf{x}) + \nabla f(\mathbf{x}+\theta\mathbf{v})\cdot\mathbf{v}$$

for some $\theta \in (0,1)$

proof: Reduce to 1-dimension case
$g(t) := f(\mathbf{x}+t\mathbf{v}), t \in \mathbb{R}$

$$\begin{aligned}
g'(t) &= \frac{d}{dt}f(\mathbf{x}+t\mathbf{v}) \\
&= \sum_{i=1}^{n}\frac{\partial f}{\partial \mathbf{x}_i}(\mathbf{x}+t\mathbf{v})\cdot\frac{d(\mathbf{x}+t\mathbf{v})_i}{dt} && \text{(by Chain Rule)} \\
&= \sum\frac{\partial f}{\partial \mathbf{x}_i}(\mathbf{x}+t\mathbf{v})\cdot\frac{d(\mathbf{x}_i+t\mathbf{v}_i)}{dt} \\
&= \sum\frac{\partial f}{\partial \mathbf{x}_i}(\mathbf{x}+t\mathbf{v})\cdot\mathbf{v}_i \\
&= \nabla f(\mathbf{x}+t\mathbf{v})\cdot\mathbf{v} && (*)
\end{aligned}$$

$g \in C^1$ on $\mathbb{R}$
Using MVT in $\mathbb{R}$:

$$\begin{aligned} f(\mathbf{x}+\mathbf{v}) &= g(1) \\ &= g(0+1) \\ &= g(0) + 1g'(0+\theta 1) && (\theta \in (0,1)) \\ &= g(0) + g'(\theta) \\ &= f(\mathbf{x}) + \nabla f(\mathbf{x}+\theta\mathbf{v}) \cdot \mathbf{v} && \text{(by (*))} \end{aligned}$$

■

2.6 1st Order Taylor Approximation in $\mathbb{R}^n$

$f \in C^1$ on $\mathbb{R}^n$

$$f(\mathbf{x}+\mathbf{v}) = f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot \mathbf{v} + o(||\mathbf{v}||)$$

proof:

$$\begin{aligned} \lim_{||\mathbf{v}||\to 0} \frac{[f(\mathbf{x}+\mathbf{v}) - f(\mathbf{x})] - \nabla f(\mathbf{x}) \cdot \mathbf{v}}{||\mathbf{v}||} &= \lim_{||\mathbf{v}||\to 0} \frac{[\nabla f(\mathbf{x}+\theta\mathbf{v}) \cdot \mathbf{v}] - \nabla f(\mathbf{x}) \cdot \mathbf{v}}{||\mathbf{v}||} \\ &= \lim_{||\mathbf{v}||\to 0} [\nabla f(\mathbf{x}+\theta\mathbf{v}) - \nabla f(\mathbf{x})] \cdot \frac{\mathbf{v}}{||\mathbf{v}||} \\ &= 0 && (\tfrac{\mathbf{v}}{||\mathbf{v}||} \text{ is a unit vector, remains 1}) \end{aligned}$$

■

2.7 2nd Order Mean Value Theorem in $\mathbb{R}^n$

$f \in C^2$ on $\mathbb{R}^n$

$$f(\mathbf{x}+\mathbf{v}) = f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot \mathbf{v} + \frac{1}{2}\mathbf{v}^T \nabla^2 f(\mathbf{x}+\theta\mathbf{v}) \cdot \mathbf{v}$$

Remarks In this course, ∇^2 means Hessian, not Laplacian.

$$\nabla^2 f(\mathbf{x}) = \left(\frac{\partial^2 f}{\partial \mathbf{x}_i \partial \mathbf{x}_j}\right)_{1\le i,j\le n} (\mathbf{x}) = \begin{pmatrix} \frac{\partial f}{\partial_1^2} & \frac{\partial f}{\partial_1 \partial_2} & \cdots \\ \frac{\partial f}{\partial_2 \partial_1} & \cdots & \\ \vdots & & \end{pmatrix}$$

The Hessian matrix is symmetric. This is sometimes called Clairaut's Theorem.
note: $\mathbf{v}^T \nabla^2 f(\mathbf{x}) \mathbf{v} = \sum_{1\le i,j\le n} \frac{\partial^2 f}{\partial \mathbf{x}_i \partial \mathbf{x}_j} f(\mathbf{x}) \mathbf{v}_i \mathbf{v}_j$

2.8 2nd Order Taylor Approximation in $\mathbb{R}^n$

$f \in C^2$ on $\mathbb{R}^n$

$$f(\mathbf{x}+\mathbf{v}) = f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot \mathbf{v} + \frac{1}{2}\mathbf{v}^T \nabla^2 f(\mathbf{x}) \mathbf{v} + o(||\mathbf{v}||^2)$$

proof:

$$\begin{aligned}\lim_{||\mathbf{v}||\to 0}\frac{[f(\mathbf{x}+\mathbf{v})-f(\mathbf{x})]-\nabla f(\mathbf{x})\cdot\mathbf{v}-\frac{1}{2}\mathbf{v}^T\nabla^2 f(\mathbf{x})\mathbf{v}}{||\mathbf{v}||^2} &= \lim_{||\mathbf{v}||\to 0}\frac{[\frac{1}{2}\mathbf{v}^T\nabla^2 f(\mathbf{x}+\theta\mathbf{v})\cdot\mathbf{v}]-\frac{1}{2}\mathbf{v}^T\nabla^2 f(\mathbf{x})\cdot\mathbf{v}}{||\mathbf{v}||^2} \\ & \qquad \text{(By 2nd MVT)} \\ &= \lim_{||\mathbf{v}||\to 0}\frac{1}{2}(\frac{\mathbf{v}}{||\mathbf{v}||})^T[\nabla^2 f(\mathbf{x}+\theta\mathbf{v})-\nabla^2 f(\mathbf{x})](\frac{\mathbf{v}}{||\mathbf{v}||}) \\ &= 0\end{aligned}$$

■

2.9 Geometric Meaning of Gradient

$f:\mathbb{R}^n\to\mathbb{R}$
Rate of change of f at $\mathbf{x}$ in direction $\mathbf{v}$ $(||\mathbf{v}||=1)=\frac{d}{dt}|_{t=0}f(\mathbf{x}+t\mathbf{v})$

$$\begin{aligned}\frac{d}{dt}|_{t=0}f(\mathbf{x}+t\mathbf{v}) &= \nabla f(\mathbf{x}+t\mathbf{v})\cdot\mathbf{v}|_{t=0} \\ &= \nabla f(\mathbf{x})\cdot\mathbf{v} \\ &= |\nabla f(\mathbf{x})||\mathbf{v}|\cos\theta \\ &= |\nabla f(\mathbf{x})|\cos\theta & (||\mathbf{v}||=1)\end{aligned}$$

maximized at $\theta=0$
So $\nabla f(\mathbf{x})$ points in the direction of steepest ascent.

2.10 Implicit Function Theorem

$f:\mathbb{R}^{n+1}\to\mathbb{R}\in C^1$
Fix $(\mathbf{a},b)\in\mathbb{R}^n\times\mathbb{R}$ s.t. $f(\mathbf{a},b)=0$.
If $\nabla f(\mathbf{a},b)\neq 0$, then $\{(\mathbf{x},y)\in(\mathbb{R}^n\times\mathbb{R})|f(\mathbf{x},y)=0\}$ is locally (near $(\mathbf{a},b)$) the graph of a function.

2.11 Level Sets of f

c-level set of $f:=\{\mathbf{x}\in\mathbb{R}^n|f(\mathbf{x})=c\}$

Fact gradient $\nabla f(\mathbf{x}_0)\perp$ level curve (through $\mathbf{x}_0$)

3 Convex Sets & Functions

3.1 Definitions

Definition of Convex Set $\Omega\subseteq\mathbb{R}^n$ is a <u>convex set</u> if $\mathbf{x}_1,\mathbf{x}_2\in\Omega\Rightarrow s\mathbf{x}_1+(1-s)\mathbf{x}_2\in\Omega$ where $s\in[0,1]$

Definition of Convex Function A function f : convex $\Omega\subseteq\mathbb{R}^n$ is <u>convex</u> if

$$f(s\mathbf{x}_1+(1-s)\mathbf{x}_2)\leq sf(\mathbf{x}_1)+(1-s)f(\mathbf{x}_2)$$

for all $\mathbf{x}_1,\mathbf{x}_2\in\Omega$ and all $s\in[0,1]$

Remarks Second line above (or equal to) the graph

Definition of Concave Function A function f is <u>concave</u> if $-f$ is convex.

3.2 Basic Properties of Convex Functions

Let $\Omega \subseteq \mathbb{R}^n$ be a convex set.

1. f_1, f_2 are convex functions on $\Omega \Rightarrow f_1 + f_2$ is a convex function on Ω.
2. f is a convex function, $a \geq 0 \Rightarrow af$ is a convex function.
3. f is a convex function on $\Omega \Rightarrow$ The sublevel sets of f, $SL_c := \{\mathbf{x} \in \mathbb{R}^n | f(\mathbf{x}) \leq c\}$ is convex.

proof of (3):
Let $x_1, x_2 \in SL_C$, so that $f(x_1) \leq c$ and $f(x_2) \leq c$.
WTS: $sx_1 + (1-s)x_2 \in SL_c$ for any $s \in [0,1]$

$$\begin{aligned} f(sx_1 + (1-s)x_2) &\leq sf(x_1) + (1-s)f(x_2) && (f \text{ is convex}) \\ &\leq sc + (1-s)c \\ &= c \\ \Rightarrow sx_1 + (1-s)x_2 \in SL_c \end{aligned}$$

■

Example of a convex function Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$
Let $x_1, x_2 \in \mathbb{R}, s \in [0,1]$
Then

$$\begin{aligned} f(sx_1 + (1-s)x_2) &= |sx_1 + (1-s)x_2| \\ &\leq |sx_1| + |(1-s)x_2| && \text{(by Triangle Inequality)} \\ &= s|x_1| + (1-s)|x_2| \\ &= sf(x_1) + (1-s)f(x_2) \end{aligned}$$

Then f is a convex function.

Theorem - Characterization of C^1 convex functions Let f : convex subset of $\mathbb{R}^n$ $\Omega \to \mathbb{R}$ be a C^1 function.
Then,
f is convex $\iff f(y) \geq f(x) + \nabla f(x) \cdot (y - x)$ for all $x, y \in \Omega$

Remarks Tangent line below the graph.

proof:
$(\Rightarrow)$
f is convex, then by definition,

$$\begin{aligned} f(s\mathbf{x}_1 + (1-s)\mathbf{x}_2) &\leq sf(\mathbf{x}_1) + (1-s)f(\mathbf{x}_2) \\ f(s\mathbf{x}_1 + (1-s)\mathbf{x}_2) - f(\mathbf{x}_2) &\leq s(f(\mathbf{x}_1) - f(\mathbf{x}_2)) \\ \frac{f(s\mathbf{x}_1 + (1-s)\mathbf{x}_2) - f(\mathbf{x}_2)}{s} &\leq f(\mathbf{x}_1) - f(\mathbf{x}_2) \\ \lim_{s\to 0} \frac{f(\mathbf{x}_2 + s(\mathbf{x}_1 - \mathbf{x}_2)) - f(\mathbf{x}_2)}{s} &\leq f(\mathbf{x}_1) - f(\mathbf{x}_2) \\ \nabla f(\mathbf{x}_2) \cdot (\mathbf{x}_1 - \mathbf{x}_2) &\leq f(\mathbf{x}_1) - f(\mathbf{x}_2) && (\text{since } \tfrac{d}{ds}|_{s=0} f(\mathbf{x}_2 + s(\mathbf{x}_1 - \mathbf{x}_2)) = \nabla f(\mathbf{x}_2) \cdot (\mathbf{x}_1 - \mathbf{x}_2)) \\ f(\mathbf{x}_2) + \nabla f(\mathbf{x}_2) \cdot (\mathbf{x}_1 - \mathbf{x}_2) &\leq f(\mathbf{x}_1) \\ f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) &\leq f(\mathbf{y}) \end{aligned}$$

where $0 \leq s \leq 1$
($\Leftarrow$)
Fix $\mathbf{x}_0, \mathbf{x}_1 \in \Omega$ and $s \in (0,1)$
Let $x = s\mathbf{x}_0 + (1-s)\mathbf{x}_1$

$$\begin{cases} f(\mathbf{x}_0) & \geq f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot (\mathbf{x}_0 - \mathbf{x}) \\ & = f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot (1-s)(\mathbf{x}_0 - \mathbf{x}_1) \\ f(\mathbf{x}_1) & \geq f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot (\mathbf{x}_1 - \mathbf{x}) \\ & = f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot s(\mathbf{x}_1 - \mathbf{x}_0) \end{cases}$$

$$\begin{cases} sf(x_0) & \geq sf(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot (1-s) \cdot s(\mathbf{x}_0 - \mathbf{x}_1) \\ (1-s)f(\mathbf{x}_1) & \geq (1-s)f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot (1-s) \cdot s(\mathbf{x}_1 - \mathbf{x}_0) \end{cases}$$

Then

$$sf(\mathbf{x}_0) + (1-s)f(\mathbf{x}_1) \geq f(x) + 0$$

Then f is convex. ■

3.3 Criterions for convexity

C^1 criterion for convexity

$$f : \Omega \to \mathbb{R} \text{ is convex } \iff f(y) \geq f(x) + \nabla f(x) \cdot (y - x)$$

for all $x, y \in \Omega$

Theorem: C^2 criterion for convexity Let $f \in C^2$ on $\Omega \subseteq \mathbb{R}^n$ (here we assume $\Omega \subseteq \mathbb{R}^n$ is a convex set containing an interior point)
Then

$$f \text{ is convex on } \Omega \iff \nabla^2 f(x) \geq 0$$

for all $x \in \Omega$

Remark 1 Let A be an $n \times n$ matrix.
"$A \geq 0$" means A is positive semi-definite:

$$v^T A v \geq 0$$

for all $v \in \mathbb{R}^n$

Remark 2 In $\mathbb{R}$,

$$f \text{ is convex } \iff f'(x) \geq 0$$

for all $x \in \Omega$
("concave up" in first year calculus)

proof for Theorem:
Recall 2nd order MVT:

$$f(y) = f(x) + \nabla f(x) \cdot (y-x) + \frac{1}{2}(y-x)^T \nabla^2 f(x + s(y-x)) \cdot (y-x)$$

for some $s \in [0,1]$
($\Leftarrow$)
Since $\nabla^2 f(x) \geq 0$, then

$$\frac{1}{2}(y-x)^T \nabla^2 f(x + s(y-x)) \cdot (y-x) \geq 0$$

Then

$$f(y) \geq f(x) + \nabla f(x) \cdot (y - x)$$

for all $x, y \in \Omega$.
Then by C^1 criterion, f is convex.
($\Rightarrow$)
Assume f is convex on Ω.
Suppose for contradiction that $\nabla^2 f(x)$ is not positive semi-definite at some $x \in \Omega$.
Then $\exists v \neq 0$ s.t. $v^T \nabla^2 f(x) v < 0$ v could be arbitrarily small and > 0
Let $y = x + v$, then

$$(y - x)^T \nabla^2 f(x + s(y - x)) \cdot (y - x) < 0$$

for all $s \in [0, 1]$
Then by MVT,

$$f(y) < f(x) + \nabla f(x) \cdot (y - x)$$

for some $x, y \in \Omega$, and this contradicts the C^1 criterion. ■

3.4 Minimization and Maximization of Convex Functions

Theorem f : convex $\Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ is a convex function.
Suppose $\Gamma := \{x \in \Omega | f(x) = \min\limits_{\Omega} f(x)\} \neq \emptyset$
(i.e. minimizer exists)
Then Γ is a convex set, and any local minimum of f is a global minimum of f.
proof:

Let $m = \min\limits_{\Omega} f(x)$.

$$\Gamma = \{x \in \Omega | f(x) = m\} = \{x \in \Omega | f(x) \leq m\}$$

(sublevel set)
Then by Basic Properties of Convex Sets, Γ is convex.
Let x be a local minimum of f.
Suppose for contradiction that $\exists y$ s.t. $f(y) < f(x)$
(i.e. x is not a global minimum)

$$\begin{aligned} f(sy + (1 - s)x) &\leq sf(y) + (1 - s)f(x) \\ &< sf(x) + (1 - s)f(x) && (f(y) < f(x)) \\ &= f(x) \end{aligned}$$

for all $s \in (0, 1)$
As s approaches 0, s approaches x.
Then we have $\lim\limits_{s \to 0} f(sy + (1 - s)x) = f(x) < f(x)$.
which is a contradiction. ■

Theorem If $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ is a convex function, and Ω is convex and compact, then

$$\max_{\Omega} f = \max_{\partial\Omega} f$$

Remarks Maximum value of f is attained (also) on the boundary of Ω
proof:
Since Ω is closed, $\partial\Omega \subseteq \Omega$, so $\max\limits_{\Omega} f \geq \max\limits_{\partial\Omega} f$.
Suppose $f(x_0) = \max\limits_{\Omega} f$ for some $x_0 \notin \partial\Omega$. Let L be an arbitrary line through x_0.

By convexity and compactness of Ω, L meets $\partial\Omega$ at two points x_1, x_2.
Let $x_0 + sx_1 + (1-s)x_2$ for $s \in (0,1)$

$$\begin{aligned} f(x_0) &= f(sx_1 + (1-s)x_2) \\ &\leq sf(x_1) + (1-s)f(x_2) && (f \text{ convex}) \\ &\leq \max\{f(x_1), f(x_2)\} \\ &\leq \max_{\partial\Omega} f \\ &\leq \max_{\Omega} f = f(x_0) \end{aligned}$$

This implies that

$$\max_{\Omega} f = \max_{\partial\Omega} f$$

as wanted. ■

Example

$$|ab| \leq \frac{1}{p}|a|^p + \frac{1}{q}|b|^q$$

where $p, q > 1$ s.t. $\frac{1}{p} + \frac{1}{q} = 1$.
Special cases:

1. $$p = q = 2, |ab| \leq \frac{|a|^2 + |b|^2}{2}$$

2. $$p = 3, q = \frac{3}{2}, |ab| \leq \frac{1}{3}|a|^3 + \frac{2}{3}|b|^{\frac{3}{2}}$$

proof:
Since function $f(x) = -\log(x)$ is convex, then

$$\begin{aligned} (-\log)|ab| &= (-\log)|a| + (-\log)|b| \\ &= \frac{1}{p}(-\log)|a|^p + \frac{1}{q}(-\log)|b|^q \\ &\geq (-\log)(\frac{1}{p}|a|^p + \frac{1}{q}|b|^q) \\ (-\log)|ab| &\geq (-\log)(\frac{1}{p}|a|^p + \frac{1}{q}|b|^q) \\ \log|ab| &\leq \log(\frac{1}{p}|a|^p + \frac{1}{q}|b|^q) \\ |ab| &\leq \frac{1}{p}|a|^p + \frac{1}{q}|b|^q && \text{(exponential function is increasing)} \end{aligned}$$

■

4 Basics of Unconstrained Optimization

4.1 Extreme Value Theorem

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is continuous, and compact set $K \subseteq \mathbb{R}^n$ Then the problem

$$\min_{x \in K} f(x)$$

has a solution.

Recall

1. $$K \subseteq \mathbb{R}^n \text{ compact} \iff K \text{ closed and bounded}$$

2. If $h_1, \dots, h_k$ and $g_1, \dots, g_m$ are continuous functions on $\mathbb{R}^n$, then the set of all points $x \in \mathbb{R}^n$ s.t.
$$\begin{cases} h_i(x) = 0 & \text{for all } i \\ g_j(x) \leq 0 & \text{for all } j \end{cases}$$
is a closed set.

3. If such a set is also bounded, then it is compact.

Example
$$\{(x, y) \in \mathbb{R}^2 | x^2 + y^2 - 1 = 0\}$$
by (2), this is a closed set
by (3), this is a compact set.

Remarks $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ convex does not imply f is continuous.

4.2 Unconstrained Optimization

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

typically

1. $\Omega \subseteq \mathbb{R}^n$
2. $\Omega = \mathbb{R}^n$
3. $\Omega =$ open
4. $\Omega = \overline{\text{open}}$

Remark

1. $\max f(x) = -(\min -f(x))$
2. $\min f(x) = -(\max -f(x))$

Definition: local minimum We say that f has a local minimum at a point $x_0 \in \Omega$ if
$$f(x_0) \leq f(x)$$
for all $x \in B_\Omega^\varepsilon(x_0)$, where $B_\Omega^\varepsilon(x_0) = \{x \in \Omega : |x - x_0| < \varepsilon\}$ which is an open ball around x_0 inside Ω of radius $\varepsilon > 0$.
We say that f has a strict local minimum at a point $x_0 \in \Omega$ if
$$f(x_0) < f(x)$$
for all $x \in B_\Omega^\varepsilon(x_0) \setminus \{x_0\}$

4.3 1st order necessary condition for local minimum

Theorem Let f be a C^1 function on $\Omega \subseteq \mathbb{R}^n$. If $x_0 \in \Omega$ is a local minimum of f, then

$$\nabla f(x_0) \cdot v \geq 0$$

for all feasible directions v at x_0

Definition: feasible direction $v \in \mathbb{R}^n$ is a feasible direction at $x_0 \in \Omega$ if

$$x_0 + sv \in \Omega$$

for all $0 \leq s \leq \bar{s}$ where $\bar{s} \in \mathbb{R}$

Remarks Feasible directions go into the set.

Corollary Special case: If $\Omega = \mathbb{R}^n$ is an open set, then any direction is a feasible direction. Then x_0 is a local minimum of f on Ω implies that $\nabla f(x_0) \cdot v \geq 0$ for all $v \in \mathbb{R}^n$.

$$\begin{cases} \nabla f(x_0) \cdot v \geq 0 \\ \nabla f(x_0) \cdot (-v) \geq 0 \iff \nabla f(x_0) \cdot v \leq 0 \end{cases} \implies \nabla f(x_0) \cdot v = 0 \text{ for all } v \in \mathbb{R}^n$$

$$\implies \nabla f(x_0) = 0$$

. ■

4.4 2nd order necessary condition for local minimum

$f \in C^2, \Omega \subseteq \mathbb{R}^n$
If $x_0 \in \Omega$ is a local minimum of f on Ω, then

1. $\nabla f(x_0) \cdot v \geq 0$ for al feasible directions v at x_0
2. If $\nabla f(x_0) \cdot v = 0$, then $v^T \nabla^2 f(x_0) v \geq 0$ (function curves up)

. ■

Remark If x_0 is an interior point of Ω, then

$$\nabla f(x_0) = 0, \quad \nabla^2 f(x_0) \geq 0$$

$$f'(x_0) = 0, \quad f''(x_0) \geq 0$$

Definition: principal minor Let A be an $n \times n$ matrix. A $k \times k$ submatrix of A formed by deleting $n-k$ rows of A, and the same $n-k$ columns of A, is called principal submatrix of A. The determinant of a principal submatrix of A is called a principal minor of A.

Definition: leading principal minor Let A be an $n \times n$ matrix. The kth order principal submatrix of A obtained by deleting the last $n-k$ rows and columns of A is called the k-th order leading principal submatrix of A, and its determinant is called a leading principal minor of A.

Definition: positive definiteness (Sylvester's Criterion) A $n \times n$ matrix A is

1. positive definite if $v^T A v > 0$ for all $v \neq 0 \iff$ all eigenvalues $> 0 \iff$ all leading principle minors > 0
2. positive semi-definite if $v^T A v \geq 0$ for all $v \iff$ all eigenvalues $\geq 0 \iff$ all principle minors ≥ 0

Lemma Suppose $\nabla^2 f(x_0)$ is positive definite, then

$$\exists a > 0 \ s.t. \ v^T \nabla^2 f(x_0) v \geq a||v||^2 \quad \forall v$$

4.5 2nd order sufficient condition (for interior points)

$f \in C^2$ on Ω
If $\begin{cases} \nabla f(x_0) = 0 \\ \nabla^2 f(x_0) > 0 \end{cases}$, then x_0 is a strict local minimum.

. ■

5 Optimization with Equality Constraints

5.1 Definitions of Related Spaces

Definition 5.1.1: surface

$$M = \text{“surface”} = \{x \in \mathbb{R}^n | h_1(x) = 0, \ldots, h_k(x) = 0\}$$

where $h_i \in C^1$

Definition 5.1.2: differentiable curve on surface A differentiable curve on surface $M \subseteq \mathbb{R}^n$ is a C^1 function

$$x : (-\epsilon, \epsilon) \to M : s \mapsto x(s)$$

Remarks

1. Let $x(s)$ be a differentiable curve on M that passes through $x_0 \in M$, say $x(0) = x_0$. The vector $v = \frac{d}{ds}|_{s=0} x(0)$ touches M “tangentially”. We say v is generated by $x(s)$.
2. In previous calculus courses, differentiable curves are often referred to as parameterizations.

Definition 5.1.3: tangent vector Any vector v which is generated by some differentiable curve on M through x_0 is called a tangent vector.

Definition 5.1.4: tangent space Tangent space to the surface M at point x_0 is

$$T_{x_0} M = \{\text{all tangent vectors to } M \text{ at } x_0\} = \{v \in \mathbb{R}^n : v = \frac{d}{ds}|_{s=0} x(s)\}$$

where $x(s)$ is a differentiable curve on M s.t. $x(0) = x_0$

Remarks The zero vector is contained in all tangent spaces.

Definition 5.1.5: T-space

$$T_{x_0} = \{x \in \mathbb{R}^n : x^T \nabla h_i(x_0) = 0 \, \forall i\} = Span\{\nabla h_1(x_0), \ldots, \nabla h_k(x_0)\}^{\perp}$$

Definition 5.1.6: regular point $x_0 \in M$ is a regular point (of the constraints) if $\{\nabla h_1(x_0), \ldots . \nabla h_k(x_0)\}$ are linearly independent.

Remark If there is only one constraint h, then x_0 is regular if and only if $\nabla h(x_0) \neq 0$.

When does the T-space equivalent to the tangent space? When x_0 is a regular point (of the constraints).

Theorem 5.1.7 Suppose x_0 is a regular point s.t. $M = \{x \in \mathbb{R}^n : h_i(x) = 0 \, \forall i\}$. Then

$$T_{x_0}M = T_{x_0}$$

Lemma 5.1.8 $f, h_1, \dots, h_k \in C^1$ on open $\Omega \subseteq \mathbb{R}^n$
$M = \{x \in \mathbb{R}^n : h_i(x) = 0 \, \forall i\}$
Suppose $x_0 \in M$ is a local minimum of f on M, then

$$\nabla f(x_0) \perp T_{x_0}M \iff \nabla f(x_0) \cdot v = 0$$

for all $v \in T_{x_0}M$

5.2 Lagrange Multipliers: 1st order necessary condition for local minimum

$f, h_1, \dots, h_k \in C^1$ on open $\Omega \subseteq \mathbb{R}^n$.
Let x_0 be a regular point of the constraints $M = \{x \in \mathbb{R}^n : h_i(x) = 0 \, \forall i\}$.
Suppose x_0 is a local minimum of f on M, then $\exists \lambda_1, \dots, \lambda_k \in \mathbb{R}$ s.t.

$$\nabla f(x_0) + \lambda_1 \nabla h_1(x_0) + \ldots + \lambda_k \nabla h_k(x_0) = 0$$

Proof. x_0 regular implies that

$$T_{x_0}M = T_{x_0} = Span\{\nabla h_1(x_0), \dots, \nabla h_k(x_0)\}^\perp$$

By Lemma 5.1.8, x_0 is a loc min implies that

$$\nabla f(x_0) \perp T_{x_0}M$$

Then

$$\nabla f(x_0) \in (T_{x_0}M)^\perp = Span\{\nabla h_i(x_0)\}^{\perp\perp} = Span\{\nabla h_i(x_0)\}$$

Then

$$\nabla f(x_0) = -\lambda_1 \nabla h_1(x_0) - \ldots - \lambda_k \nabla h_k(x_0)$$

for some $\lambda_i \in \mathbb{R}$ ■

5.3 2nd order necessary condition for local minimum

$f, h_1, \dots, h_k \in C^2$ on open $\Omega \subseteq \mathbb{R}^n$.
Let x_0 be a regular point of the constraints $M = \{x \in \mathbb{R}^n : h_i(x) = 0 \, \forall i\}$.
Suppose x_0 is a local minimum of f on M, then

1. $$\nabla f(x_0) + \sum_{i=1}^{k} \lambda_i \nabla h_i(x_0) = 0$$

   for some $\lambda_i \in \mathbb{R}$

2. $$\nabla^2 f(x_0) + \sum \lambda_i \nabla^2 h_i(x_0) \geq 0$$

   on $T_{x_0}M$

5.4 2nd order sufficient condition for local minimum

$f, h_1, \ldots, h_k \in C^2$ on open $\Omega \subseteq \mathbb{R}^n$.
Let x_0 be a regular point of the constraints $M = \{x \in \mathbb{R}^n : h_i(x) = 0 \, \forall i\}$.
If $\exists \lambda_i \in \mathbb{R}$ s.t.

1. $$\nabla f(x_0) + \sum \lambda_i \nabla h_i(x_0) = 0$$

2. $$\nabla^2 f(x_0) + \sum \lambda_i \nabla^2 h_i(x_0) > 0$$

   on $T_{x_0}M$

Then x_0 is a strict local minimum.

Proof. Recall that (2) means $[\nabla^2 f(x_0) + \sum \lambda_i h_i(x_0)]$ is pos-def on $T_{x_0}M$.
Then $\exists a > 0$ s.t. $v^T[\nabla^2 f(x_0) + \sum \lambda_i h_i(x_0)]v \geq a||v||^2$ for all $v \in T_{x_0}M$.
Let $x(s) \in M$ be a curve s.t. $x(0) = x_0$ and $v = x'(0)$.
WLOG, $||x'(0)|| = 1$.
By 2nd order Taylor,

$$\begin{aligned}
f(x(s)) - f(x(0)) &= s\frac{d}{ds}|_{s=0} f(x(s)) + \frac{1}{2}s^2\frac{d}{ds^2}|_{s=0} f(x(s)) + o(s^2) \\
&= s\frac{d}{ds}|_{s=0}[f(x(s)) + \sum_i \lambda_i h_i(x(s))] + \frac{1}{2}s^2\frac{d}{ds^2}|_{s=0}[f(x(s)) + \sum_i \lambda_i h_i(x(s))] + o(s^2) \\
& \qquad (\textstyle\sum_i \lambda_i h_i(x(s)) = 0) \\
&= s[\nabla f(x_0) + \sum \lambda_i \nabla h_i(x_0)] \cdot x'(0) + \frac{1}{2}s^2 x'(0)^T[\nabla^2 f(x_0) + \sum \lambda_i \nabla^2 h_i(x_0)]x'(0) + o(s^2) \\
&= 0 + \frac{1}{2}s^2 v^T[\nabla^2 f(x_0) + \sum \lambda_i \nabla^2 h_i(x_0)]v + o(s^2) \\
&\geq \frac{1}{2}s^2 a||v||^2 + o(s^2) \\
&= \frac{1}{2}s^2 a + o(s^2) \\
&= s^2[\frac{a}{2} + \frac{o(s^2)}{s^2}] > 0
\end{aligned}$$

for small $s > 0$, since $\frac{a}{2} > 0$ and $\lim_{s \to 0} \frac{o(s^2)}{s^2} = 0$
Then $f(x(s)) > f(x_0)$ for small $s > 0$ Then x_0 is a strict local min of f. ∎

6 Optimization with Inequality Constraints

Problem open $\Omega \subseteq \mathbb{R}^n$
$f : \Omega \to \mathbb{R}$
$h_1, \ldots, h_k : \Omega \to \mathbb{R}$
$g_1, \ldots, g_l : \Omega \to \mathbb{R}$

$$\begin{cases} \min f(x) \\ x \in \Omega \text{ subject to } \begin{cases} h_1(x) = 0, \ldots, h_k(x) = 0 \\ g_1(x) \leq 0, \ldots, g_l(x) \leq 0 \end{cases} \end{cases} \qquad (*)$$

Definition 1: activeness Let x_0 satisfy the constraints.
We say that the constraint $g_i(x) \leq 0$ is active at x_0 if $g_i(x_0) = 0$.
It is inactive at x_0 if $g_i(x_0) < 0$.

Definition 2: regular point Suppose for some $l' \leq l$:

$$g_1(x) \leq 0, \dots, g_{l'}(x) \leq 0;\ g_{l'+1}(x) \leq 0, \dots, g_l(x) \leq 0$$

where $g_1, \dots g_{l'}$ active and the rest inactive.
We say that x_0 is a regular point of the constraints if
$\{\nabla h_1(x_0), \dots, \nabla h_k(x_0), \nabla g_1(x_0), \dots, \nabla g_{l'}(x_0)\}$ is linearly independent.

6.1 Kuhn-Tucker conditions: 1st order necessary condition for local minimum

open $\Omega \subseteq \mathbb{R}^n$
$f : \Omega \to \mathbb{R}$
$h_1, \dots, h_k, g_1, \dots, g_l : C^1 \in \Omega$
Suppose $x_0 \in \Omega$ is a regular point of the constraints which is a local minimum, then

1. $$\nabla f(x_0) + \sum_{i=1}^{k} \lambda_i \nabla h_i(x_0) + \sum_{j=1}^{l} \mu_j \nabla g_j(x_0) = 0$$

   for some $\lambda_i \in \mathbb{R}$ and $\mu_j \geq 0$

2. $\mu_j g_j(x_0) = 0$

Remark 1 Given x_0,

$$\begin{cases} g_j(x) \leq 0 \text{ active at } x_0 \implies g_j(x_0) = 0 \implies \mu_j g_j(x_0) = 0 \\ g_j(x) \leq 0 \text{ inactive at } x_0 \implies g_j(x_0) < 0 \implies \mu_j = 0 \end{cases}$$

$\implies \mu_j = 0$ for all inactive g_j at x_0

Remark 2 It is possible for an active constraint to have zero multiplier.

Remark 3 $\mu_j \geq 0$ because ∇f and ∇g have opposite directions at a local minimum x_0.

$$\nabla f(x_0) + \mu \nabla g(x_0) = 0 \implies \nabla f(x_0) = -\mu \nabla g(x_0) \implies -\mu < 0 \implies \mu > 0$$

Is this true?

Idea of proof x_0 is a local min of f subject to (*)
$\implies$ x_0 is a local min for equality constraints $h_1(x) = 0, \dots, h_k(x) = 0+$ active inequality constraints $g_1(x) \leq 0, \dots, g_{l'}(x) \leq 0$
$\implies$ x_0 is a local min for $h_1(x) = 0, \dots, h_k(x) = 0 + g_1(x) = 0, \dots, g_{l'}(x) = 0 \implies \nabla f(x_0) + \sum_{i=1}^{k} \lambda_i \nabla h_i(x_0) + \sum_{j=1}^{l'} \mu_j \nabla g_j(x_0) = 0$
for some $\lambda_i \in \mathbb{R}$ and $\mu_j \in \mathbb{R}$.
Let $\mu_j = 0$ for $j = l' + 1, \dots, l$, then

$$\nabla f(x_0) + \sum_{i=1}^{k} \lambda_i \nabla h_i(x_0) + \sum_{j=1}^{l} \mu_j \nabla g_j(x_0) = 0$$

6.2 2nd order necessary conditions for local minimum

Open $\Omega \subseteq \mathbb{R}^n, f, h_1, \ldots, h_k, g_1, \ldots, g_l \in C^2$. Let x_0 be a regular point of the constraints:

$$(\dagger) \begin{cases} h_1(x) = \ldots = h_k(x_0) = 0 \\ g_1(x), \ldots g_l(x_0) \leq 0 \end{cases}$$

Suppose x_0 is a local min of f subject to ($\dagger$). Then, $\exists \lambda_i \in \mathbb{R}, \mu_j \geq 0$ s.t.

1. $\nabla f(x_0) + \sum_{i=1}^{k} \lambda_i \nabla h_i(x_0) + \sum_{j=1}^{l} \mu_j \nabla g_j(x_0) = 0$
2. $\mu_j g_j(x_0) = 0$ for all j
3. $[\nabla^2 f(x_0) + \sum \lambda_i \nabla^2 h_i(x_0) + \sum \mu_j \nabla^2 g_h(x_0)]]$ is positive semi definite on tangent space to active constraints at x_0.

Proof. x_0 local min for ($\dagger$)
$\implies x_0$ local min for only active constraints at x_0.
$\implies \begin{cases} h_i(x) = 0 \ \forall i \\ g_j(x) = 0 \ j = 1, \ldots, l' \end{cases}$
$\implies [\nabla^2 f(x_0) + \sum \lambda_i \nabla^2 h_i(x_0) + \sum \mu_j \nabla^2 g_h(x_0)]$ pos semi def on tangent space to active constraints. ■

6.3 2nd order sufficient conditions

Open $\Omega \subseteq \mathbb{R}^n, f, h_i, g_j \in C^2$ on Ω.
Problem:

$$\begin{cases} \min & f(x) \\ \text{subject to} & \begin{cases} h_i(x) = 0 \\ g_j(x) \leq 0 \end{cases} \end{cases}$$

Suppose $\exists x_0$ feasible and $\lambda_i, \mu_j \in \mathbb{R}$ s.t.

1. $\nabla f(x_0) + \sum_{i=1}^{k} \lambda_i \nabla h_i(x) + \sum_{j=1}^{l} \mu_j \nabla g_j(x_0) = 0$
2. $\mu_j g_j(x_0) = 0$ all j

If the Hessian matrix, $L(x_0) = \nabla^2 f(x_0) + \sum_{i=1}^{k} \lambda_i \nabla^2 f(x_0) + \sum_{i=1}^{l} \mu_j \nabla^2 g_j(x)$ is pos def on $\tilde{\tilde{T}}_{x_0}$-space of "strongly active" constraints at x_0.
Then x_0 is a strict local min.

Remarks

1. $$\text{Active constraints at } x_0 \begin{cases} h_i(x) = 0 & i = 1, \ldots, k \\ g_j(x) \leq 0 & j = 1, \ldots, l' \implies g_j(x_0) = 0 \end{cases}$$

2. $$\text{Strongly active constraints at } x_0 \begin{cases} h_i(x) = 0 & i = 1, \ldots, k \\ g_j(x) \leq 0 & j = 1, \ldots, l'' \ \ g_j(x) \text{ is active at } x_0 \text{ and } \mu_j > 0 \end{cases}$$

   $l'' \leq l' \leq l$

3. $$\tilde{\tilde{T}}_{x_0} = \{v \in \mathbb{R}^n | v \cdot \nabla h_i(x_0) = 0 \text{ all } i \text{ and } v \cdot \nabla g_j(x_0) = 0 \text{ for all } j = 1, \ldots, l''\}$$

4. strongly active $\subseteq$ active
   $\implies \tilde{\tilde{T}}_{x_0} = (\text{strongly active})^\perp \supseteq (\text{active})^\perp = \tilde{T}_{x_0}$

Proof. (details see another pdf by prof) Suppose x_0 is NOT a (strict) local min.
claim: $\exists$ unit vector $v \in \mathbb{R}$ s.t.

1. $\nabla f(x_0) \cdot v \leq 0$
2. $\nabla h_i(x_0) \cdot v = 0 \quad i = 1, \ldots, k$
3. $\nabla g_j(x_0) \cdot v \leq 0 \quad j = 1, \ldots, l'$

proof of claim: []
claim: $\nabla g_j(x) \cdot v = 0$ for $j = 1, \ldots, l''$
proof of claim: []
$\implies$ contradiction!
claim: $\exists$ unit vector $v \in \mathbb{R}$ s.t.

1. $\nabla f(x_0) \cdot v \leq 0$
2. $\nabla h_i(x_0) \cdot v = 0 \quad i = 1, \ldots, k$
3. $\nabla g_j(x_0) \cdot v = 0 \quad j = 1, \ldots, l''$

proof of claim: [] ■

7 Different Computation Methods for Solving Optimum

7.1 Newton's Method

$x_0 \in I$ start
$x_{n+1} = x_0 - \frac{f'(x_0)}{f''(x_0)}$

Theorem Let $f \in C^3$ on I.
Suppose $x_* \in I$ satisfies $f'(x_*) = 0$ and $f''(x_*) \neq 0$ (x_* is a non-degenerate (non-singular) critical point).
Then the sequence of points $\{x_n\}$ generated by Newton's method

$$x_{n+1} = x_n - \frac{f'(x_0)}{f''(x_0)}$$

converges to x_* if x_0 is sufficiently close to x_*.

Why do we need this method? In real life, we may not know the real function formula. We only have data, using which we can approximate the function formula. In a way, Newton's method is true "applied mathematics".

Proof of Theorem Let $g(x) = f'(x)$ so that $x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)}$
By $g \in C^2, \exists \alpha$ s.t. $|g'(x_1)| > \alpha \, \forall x_1$ and $|g''(x_2)| < \frac{1}{\alpha} \, \forall x_2$ in a neighbourhood of x_* (choose α small enough).

$$
\begin{aligned}
x_{n+1} - x_* &= x_n - \frac{g(x_n)}{g'(x_n)} - x_* && (1)\\
&= x_n - x_* - \frac{g(x_n) - g(x_*)}{g'(x_n)} \qquad (g(x_*) = 0) && (2)\\
&= \frac{-[g(x_n) - g(x_*) - g'(x_n)(x_n - x_*)]}{g'(x_n)} && (3)\\
&= \frac{1}{2}\frac{g''(\xi)}{g'(x_n)}(x_n - x_*)^2 && (4)\\
|x_{n+1} - x_*| &= \frac{1}{2}\frac{g''(\xi)}{g'(x_n)}|x_n - x_*|^2 < \frac{1}{2\alpha^2}|x_n - x_*|^2 \qquad \text{(in small neighbourhood of } x_*) && (5)\\
\rho &:= \frac{1}{2\alpha^2}|x_0 - x_*| \qquad \text{(choose } x_0 \text{ sufficiently close to } x_* \text{ s.t. } \rho < 1) && (6)\\
|x_1 - x_*| &< \frac{1}{2\alpha^2}|x_0 - x_*|^2 && (7)\\
&= \frac{1}{2\alpha^2}|x_0 - x_*||x_0 - x_*| && (8)\\
&= \rho|x_0 - x_*| && (9)\\
|x_2 - x_*| &< \frac{1}{2\alpha^2}|x_1 - x_*|^2 && (10)\\
&< \frac{1}{2\alpha^2}\rho^2|x_0 - x_*|^2 && (11)\\
&= \frac{1}{2\alpha^2}|x_0 - x_*|\rho^2|x_0 - x_*| && (12)\\
&< \rho^2|x_0 - x_*| \qquad (\rho < 1) && (13)\\
|x_n - x_*| &< \rho^n|x_0 - x_*| \underset{n\to\infty}{\to} 0 && (14)\\
\implies & x_n \to x_* && (15)
\end{aligned}
$$

<u>proof of (4):</u>
By 2nd order MVT,

$$g(x) = g(y) + g'(y)(x - y) + \frac{1}{2}g''(\xi)(x - y)^2$$

for some $\xi \in [x, y]$.
Let $x = x_*$ and $y = x_n$, then

$$g(x_*) = g(x_n) + g'(x_n)(x_* - x_n) + \frac{1}{2}g''(\xi)(x_* - x_n)^2$$

$$\implies -[g(x_n) - g(x_*) - g'(x_n)(x_n - x_*)] = \frac{1}{2}g''(\xi)(x_n - x_*)^2$$

■

Newton's Method (generalized) $f : \Omega \underset{open}{\subseteq} \mathbb{R}^n \to \mathbb{R}$ and $f \in C^3$ on Ω

$x_0 \in \Omega$
$x_{n+1} = x_n - [\nabla^2 f(x_n)]^{-1}\nabla f(x_n)$
(The algorithm requires $\nabla^2 f(x_n)$ invertible and stops when $\nabla f(x_n) = 0$)

Note Newton's method may fail to converge even if $f(x)$ has a unique global min x_* and x_0 is arbitrarily close to x_*

Remark Newton's method, if converge, converges to

1. local min
2. local max
3. saddle point

Example 7.1. Newton's Method on Quadratic Function Let Q be a symmetric $n \times n$ invertible matrix. Define quadratic form $f(x) := \frac{1}{2}x^T Qx : \mathbb{R}^n \to \mathbb{R}$. Then the optima is $x = 0$.
Let $x_0 \in \mathbb{R}^n$, then

$$x_1 := x_0 - \nabla^2 f(x_0)^{-1} \nabla f(x_0) = x_0 - Q^{-1}Qx_0 = 0$$

Newton's method converges in one iteration.

7.2 Method of Steepest Descent (Gradient Method)

$f : \Omega \underset{open}{\subseteq} \mathbb{R}^n \to \mathbb{R}, C^1$
Recall: Direction of steepest ascent at x_0 is given by the direction of gradient $\nabla f(x_0)$

Algorithm of steepest descent $x_0 \in \Omega$

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$$

where $\alpha_k \geq 0$ satisfying $f(x_k - \alpha_k \nabla f(x_k)) = \min_{\alpha \geq 0} f(x_k - \alpha \nabla f(x_k))$
(keep going until you find the minimum)

Fact: algorithm is descending If $\nabla f(x_k) \neq 0$, then $f(x_{k+1}) < f(x_k)$
Why? $f(x_{k+1}) = f(x_k - \alpha_k \nabla f(x_k)) \leq f(x_k - \alpha \nabla f(x_k))$ for all $0 < \alpha \leq \alpha_k$
Recall: $\frac{d}{ds}|_{s=0} f(x_k - s\nabla f(x_k)) = \nabla f(x_k) \cdot (-\nabla f(x_k)) = -|\nabla f(x_k)|^2 < 0$
$\implies f(x_{k+1}) \leq f(x_k - \alpha \nabla f(x_k)) < f(x_k)$ for small α

Fact: the method of steepest descent moves perpendicular steps

$$\begin{aligned}(x_{k+2} - x_{k+1}) \cdot (x_{k+1} - x_k) &= (-\alpha_{k+1} \nabla f(x_{k+1})) \cdot (-\alpha_k \nabla f(x_k)) && (16)\\ &= \alpha_k \alpha_{k+1} \nabla f(x_{k+1}) \cdot \nabla f(x_k) && (17)\\ & && (18)\end{aligned}$$

If $\alpha_k = 0$, then we are done.
If $\alpha_k \neq 0$, then

$$\nabla f(x_{k+1}) = \min_{\alpha \geq 0} f(x_k - \alpha \nabla f(x_k)) \tag{19}$$

$$\implies \frac{d}{d\alpha}|_{\alpha=\alpha_k} f(x_k - \alpha \nabla f(x_k)) = (-\nabla f(x_k)) \cdot \nabla f(x_k - \alpha_k \nabla f(x_k)) = 0 \tag{20}$$

$$\implies \alpha_k \alpha_{k+1} \nabla f(x_{k+1}) \cdot \nabla f(x_k) = 0 \tag{21}$$

Note This method is not the most efficient. May take infinite steps to converge.

Theorem (Convergence of Steepest Descent) $f \in C^1$ on $\Omega \underset{open}{\subseteq} \mathbb{R}^n$

Let $\{x_k\}$ be sequence generated by steepest descent.
$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$
If $\{x_k\}$ is "bounded in Ω" (i.e. $\exists$ compact set $K \subset \Omega$ s.t. $x_k \in K$ for all k)
Then every convergent subsequence of $\{x_k\}$ converges to a critical point $x_* \in \Omega$ of $f : \nabla f(x_*) = 0$

Proof. $x_k \in$ compact $K \implies$ subsequence $x_{k_i} \to x_* \in K$
Since $f(x_0) \geq f(x_1) \geq f(x_2) \geq \ldots$ and $f(x_{k_i}) \searrow f(x_*)$
Suppose by contradiction that $\nabla f(x_*) \neq 0$
$x_{k_i} \to x_* \implies \nabla f(x_{k_i}) \to \nabla f(x_*)$
Let $y_{k_i} = x_{k_i} - \alpha_{k_i} \nabla f(x_{k_i}) = x_{k_i+1}$. Then $y_{k_i} \to y_*$. Then

$$f(y_{k_i}) = f(x_{k_i+1}) = \min_{\alpha \geq 0} f(x_i - \alpha \nabla f(x_{k_i})) \tag{22}$$
$$f(y_{k_i}) \leq f(x_{k_i} - \alpha \nabla f(x_{k_i})) \text{ for all } \alpha \geq 0 \tag{23}$$
$$\lim_{i \to \infty} f(y_*) \leq f(x_* - \alpha \nabla f(x_*)) \text{ for all } \alpha \geq 0 \tag{24}$$
$$f(y_*) \leq \min_{\alpha \geq 0} f(x_* - \alpha \nabla f(x_*)) < f(x_*) \tag{25}$$
$$f(y_*) < f(x_*) \tag{26}$$
$$\tag{27}$$

But $f(y_*) = \lim_{i \to \infty} f(y_{k_i}) = \lim_{i \to \infty} f(x_{k_i+1}) = f(x_*)$, so we have a contradiction. ■

Steepest descent: Quadratic case Let f follow the general quadratic form

$$f(x) = \frac{1}{2} x^T Q x - b^T x$$

where $b, x \in \mathbb{R}^n$ and Q is an $n \times n$ positive definite matrix.
Let $0 < \lambda = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n = \Lambda$ be eigenvalues of Q.
Recall that if Q pos-def, then there is a unique minimum x_* such that $Qx_* - b = 0 \iff x_* = Q^{-1} b$
Define $q(x) := \frac{1}{2}(x - x^*)^T Q (x - x^*) = f(x) + const$
Note that $q(x) \geq 0$ and $q(x_*) = 0$.
Define $g(x) := Qx - b = \nabla q(x) = \nabla f(x)$
So using the method of steepest descent:

$$x_{k+1} = x_k - \alpha_k g(x_k)$$

Derive the formula for α_k:
α_k minimizes $f(x_k - \alpha g(x_k))$

$$0 = \frac{d}{d\alpha} |_{\alpha = \alpha_k} f(x_k - \alpha g(x_k)) \tag{28}$$
$$= \nabla f(x_k - \alpha_k g(x_k)) \cdot (-g(x_k)) \tag{29}$$
$$= -[Q(x_k - \alpha_k g(x_k)) - b] \cdot (g(x_k)) \tag{30}$$
$$= -(Qx_k - b) \cdot g(x_k) \tag{31}$$
$$= -|g(x_k)|^2 + \alpha_k g(x_k)^T Q g(x_k) \tag{32}$$
$$\implies \alpha_k = \frac{|g(x_k)|^2}{g(x_k)^T Q g(x_k)} \tag{33}$$
$$\implies x_{k+1} = x_k - \alpha_k g(x_k) \tag{34}$$
$$= x_k - \frac{|g(x_k)|^2}{g(x_k)^T Q g(x_k)} g(x_k) \tag{35}$$

Claim:

$$q(x_{k+1}) = \left(1 - \frac{|g(x_k)|^4}{(g(x_k)^T Q g(x_k))(g(x_k)^T Q^{-1} g(x_k))}\right) g(x_k)$$

Proof.

$$\begin{aligned}
q(x_{k+1}) &= q(x_k - \alpha_k g(x_k)) && (36)\\
&= \frac{1}{2}(x_k - \alpha_k g(x_k) - x_*)^T Q (x_k - \alpha_k g(x_k) - x_*) && (37)\\
&= \frac{1}{2}(x_k - x_* - \alpha_k g(x_k))^T Q((x_k - x_*) - \alpha_k g(x_k)) && (38)\\
&= \frac{1}{2}(x_k - x_*)^T Q(x_k - x_*) - \alpha_k g(x_k)^T Q(x_k - x_*) + \frac{1}{2}\alpha_k^2 g(x_k)^T Q g(x_k) && (39)\\
&= q(x_k) - \alpha_k g(x_k)^T Q(x_k - x_*) + \frac{1}{2}\alpha_k^2 g(x_k)^T Q g(x_k) && (40)\\
\implies q(x_k) - q(x_{k+1}) &= -\frac{1}{2}\alpha_k^2 g(x_k)^T Q g(x_k) + \alpha_k g(x_k)^T Q(x_k - x_*) && (41)\\
y_k &:= x_k - x_* && (42)\\
\frac{q(x_k) - q(x_{k+1})}{q(x_k)} &= \frac{-\frac{1}{2}\alpha_k^2 g(x_k)^T Q g(x_k) + \alpha_k g(x_k)^T Q y_k}{\frac{1}{2} y_k^T Q y_k} && (43)\\
&= \frac{2\alpha_k g(x_k)^T Q y_k - \alpha_k^2 g(x_k)^T Q g(x_k)}{y_k^T Q y_k} && (44)\\
(g_k := g(x_k) &= Q x_k - b = Q x_k - Q x_* = Q(x_k - x_*) = Q y_k \implies y_k = Q^{-1} g_k) && (45)\\
&= \frac{2\alpha_k |g_k|^2 - \alpha_k^2 g_k^T Q g_k}{g_k^T Q^{-1} g_k} && (46)\\
&= \frac{2\frac{|g_k|^4}{g_k^T Q g_k} - \frac{|g_k|^4}{g_k^T Q g_k}}{g_k^T Q^{-1} g_k} && (47)\\
&= \frac{|g_k|^4}{(g_k^T Q g_k)(g_k^T Q^{-1} g_k)} && (\alpha_k = \tfrac{|g(x_k)|^2}{g(x_k)^T Q g(x_k)})\\
\implies q(x_k) - q(x_{k+1}) &= \left(\frac{|g_k|^4}{(g_k^T Q g_k)(g_k^T Q^{-1} g_k)}\right) q(x_k) && (48)\\
\implies q(x_{k+1}) &= q(x_k)\left(1 - \frac{|g_k|^4}{(g_k^T Q g_k)(g_k^T Q^{-1} g_k)}\right) && (49)\\
&\le (1 - \frac{4\lambda\Lambda}{(\lambda + \Lambda)^2}) q(x_k) && \text{(By Kantorovich Inequality)}\\
\implies q(x_{k+1}) &\le \frac{(\Lambda - \lambda)^2}{(\lambda + \Lambda)^2} q(x_k) && (50)
\end{aligned}$$

■

Kantorovich Inequality $Q : n \times n$ positive definite symmetric matrix
$\lambda = \lambda_1 \le \lambda_2 \le \ldots \le \lambda_n = \Lambda$
For any $v \in \mathbb{R}^n$:

$$\frac{|v|^4}{(v^T Q v)(v^T Q^{-1} v)} \ge \frac{4\lambda\Lambda}{(\lambda + \Lambda)^2}$$

Theorem: Steepest Descent in Quadratic Case For any $x_0 \in \mathbb{R}^n$, method of steepest descent converges to the unique min point x_* of f.

Furthermore, for $q(x) := \frac{1}{2}(x - x_*)Q(x - x_*)$, where Q symmetric positive definite and $0 < \lambda = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n = \Lambda$,

$$q(x_{k+1}) \leq \frac{(\Lambda - \lambda)^2}{(\lambda + \Lambda)^2} q(x_k)$$

Let $r := \frac{(\Lambda-\lambda)^2}{(\lambda+\Lambda)^2}$, then

$$q(x_k) \leq r^k q(x_0)$$

for all k. As $k \to \infty$, $q(x_k) \to 0$.

Notes

1. $x_k \in \{x \in \mathbb{R}^n | q(x) \leq r^k q(x_0)\} = SL_k$
(sublevel set of function $q(x)$)
Note that SL_k is strictly decreasing. Furthermore, note that x_* is the only point satisfying the inequality at the limit:
$$q(x_*) = 0 = \lim_{k \to \infty} q(x_0)$$
Therefore, $\lim_{k \to \infty} SL_k = \{0\}$, and $x_k \to x_*$.

2. $r = (\frac{(\Lambda-\lambda)}{(\lambda+\Lambda)})^2 = (\frac{\Lambda/\lambda-1}{\Lambda/\lambda+1})^2$ depends only on the ratio $\frac{\Lambda}{\lambda}$ = "condition number of Q"
case $\frac{\Lambda}{\lambda} = 1 \implies r = 0 \implies 0 \leq q(x_1) \leq 0 \cdot q(x_0) \implies q(x_1) = 0 \implies x_1 = x_*$
(Gradient descent converges to the unique global minimum in only one iteration.)
case $\frac{\Lambda}{\lambda} >> 1 \implies r \simeq 1$
(worst case, converges very flow)

7.3 Method of Conjugate Direction

Motivation Method of conjugate directions is designed for quadratic functions with form $f(x) = \frac{1}{2}x^T Qx - b^T x$. For other functional forms, one can approximate the function using quadratic form firstly and then apply method of conjugate directions.

Definition: Q-orthogonality Let Q be a symmetric matrix. Two vectors $d, d' \in \mathbb{R}^n$ are Q-orthogonal (or Q-conjugate) if

$$d^T Qd' = 0$$

A finite set of $d_0, \ldots, d_k$ is called Q-orthogonal set if $d_i^T Qd_j = 0$ for all $i \neq j$.

Example 1 Q is an identity matrix. d, d' are Q-orthogonal iff they are orthogonal.

Example 2 If d, d' are two eigenvectors with different eigenvalues, then they are Q-orthogonal.

Proof. Suppose $Qv = \lambda v$ and $Qw = \lambda' w$ so $\lambda \neq \lambda'$

$$< v, Qw > = < v, \lambda' w > = \lambda' < v, w > \quad (51)$$
$$= < Q^T v, w > = < Qv, w > = < \lambda v, w > = \lambda < v, w > \quad (52)$$
$$\implies (\lambda - \lambda')\langle v, w\rangle = 0 \quad (53)$$

Since $(\lambda - \lambda') \neq 0$, them we have $\langle v, w\rangle = 0$.

$$\implies v^T Qw = \langle v, Qw\rangle = \lambda\langle v, w\rangle = 0$$

■

Example 3 If Q is an $n \times n$ symmetric matrix, then there exists an orthogonal basis of eigenvectors $d_0, \ldots, d_{n-1}$
Claim: They are also Q-orthogonal.

Proof. $d_i^T Q d_j = d_i^T(\lambda d_j) = \lambda d_i^T d_j = 0$ ∎

Proposition Let Q be a symmetric positive definite matrix. Let $d_0, \ldots, d_k$ be a set of (non-zero) Q-orthogonal vectors. Then $d_0, \ldots, d_k$ are linearly independent.

Proof. Assume $\alpha_0 d_0 + \ldots + \alpha_k d_k = 0$ for $\alpha_i \in \mathbb{R}$.
Multiply the whole equation by $d_i^T Q$:

$$\alpha_0 \underbrace{d_i^T Q d_0}_{=0} + \ldots + \alpha_i \underbrace{d_i^T Q d_i}_{>0} + \ldots + \alpha_k \underbrace{d_i^T Q d_k}_{=0} = 0$$

which implies $\alpha_i d_i^T Q d_i = 0$ and $\alpha_i = 0$.
This is true for every i. Therefore $d_0, \ldots, d_k$ are linearly independent. ∎

Lemma (Theorems covered so far)

1. d_i, d_j are Q-orthogonal if $d_i^T Q d_j = 0$;
2. Eigen-vectors with different eigenvalues are Q-orthogonal;
3. Matrix Q symmetric $\implies$ there exists an orthogonal basis $\implies$ the set of basis is Q-orthogonal as well;
4. Q-orthogonal vectors are linearly independent.

Example 4 (Special case: Method of Conjugate Direction on Quadratic Functions). Let Q be a positive definite symmetric $n \times n$ matrix. The problem is

$$\min f(x) = \frac{1}{2} x^T Q x - b^T x$$

Recall that the unique global minimum is $x^* = Q^{-1} b$.
Let $d_0, d_1, \ldots, d_{n-1}$ be non-zero Q-orthogonal vectors.
Note that they are linearly independent by the previous theorem.
Therefore, they form a basis of $\mathbb{R}^n$.
The global minimum can be represented as

$$x^* = \sum_{j=0}^{n-1} \alpha_j d_j,\ \alpha_j \in \mathbb{R}$$

For every j, the following holds:

$$d_j^T Q x^* = \alpha_j d_j^T Q d_j$$

$$\implies \alpha_j = \frac{d_j^T Q x^*}{d_j^T Q d_j}$$

Algorithm: Method of Conjugate Directions Let Q be a positive definite symmetric $n \times n$ matrix. and $\{d_j\}_{j=0}^{n-1}$ be a set of non-zero Q-orthogonal vectors, note that they form a basis of $\mathbb{R}^n$.
Given initial point $x_0 \in \mathbb{R}^n$, the method of conjugate direction generates a sequence of points $\{x_k\}_{k=0}^n$ as the following:

$$x_{k+1} \leftarrow x_k + \alpha_k d_k$$

$$\alpha_k := -\frac{\langle g_k, d_k \rangle}{d_k^T Q d_k} \quad g_k := \nabla f(x_k)$$

Theorem Given the method of conjugate, the sequence of points generated eventually reaches the global minimum. That is, $x_n = x^*$.

Proof. Let $x^*, x_0 \in \mathbb{R}^n$, consider

$$x^* - x_0 = \sum_{j=0}^{n-1} \beta_j d_j \tag{54}$$

$$\iff x^* = x_0 + \sum_{j=0}^{n-1} \beta_j d_j \tag{55}$$

$$d_j^T Q(x^* - x_0 = d_j^T Q(\sum_{j=0}^{n-1} \beta_j d_j) \tag{56}$$

$$= \beta_j d_j^T Q d_j \tag{57}$$

$$\implies \beta_j = \frac{d_j^T Q(x^* - x_0)}{d_j^T Q d_j} \tag{58}$$

Note that the algorithm generates the sequence as following:

$$x_k = x_0 + \sum_{j=0}^{k-1} \alpha_j d_j \tag{59}$$

$$\implies (x_k - x_0) = \sum_{j=0}^{k-1} \alpha_j d_j \tag{60}$$

$$\implies d_k^T Q(x_k - x_0) = \sum_{j=0}^{k-1} \alpha_j d_k^T Q d_j = 0 \tag{61}$$

Therefore,

$$\beta_k = \frac{d_k^T Q(x^* - x_0)}{d_k^T Q d_k} \tag{62}$$

$$= \frac{d_k^T Q(x^* - x_0) - d_k^T Q(x_k - x_0)}{d_k^T Q d_k} \tag{63}$$

$$= \frac{d_k^T Q(x^* - x_k)}{d_k^T Q d_k} \tag{64}$$

$$= \frac{d_k^T (Qx^* - Qx_k)}{d_k^T Q d_k} \tag{65}$$

$$= \frac{d_k^T (b - Qx_k)}{d_k^T Q d_k} \qquad \text{(The first order necessary condition suggests } Qx^* = b\text{)}$$

$$= -\frac{d_k^T (Qx_k - b)}{d_k^T Q d_k} \tag{66}$$

$$= -\frac{d_k^T \nabla f(x_k)}{d_k^T Q d_k} \qquad \text{(Assuming } f \text{ is quadratic)}$$

$$= \alpha_k \tag{67}$$

Consequently,

$$x^* = x_0 + \sum_{j=0}^{n-1} \beta_j d_j \tag{68}$$
$$= x_0 + \sum_{j=0}^{n-1} \alpha_j d_j \tag{69}$$
$$= x_n \tag{70}$$

■

7.3.1 Geometric Interpretations of Method of Conjugate Directions

Theorem Let $f \in C^1(\Omega, \mathbb{R})$, where Ω is a convex subset of $\mathbb{R}^n$, then x_0 is a local minimum of f on Ω if and only if

$$\nabla f(x_0) \cdot (y - x_0) \geq 0 \, \forall y \in \Omega$$

Example Now consider the special case in which Ω is an affine hyperplane, that is,

$$\Omega = \{x \in \mathbb{R}^n : cx + b = 0\}$$

where $\dim(\Omega)$ is $n-1$.
Note that for every $y \in \Omega, \nabla f(x_0) \cdot (y - x_0) \geq 0$. For any feasible direction $a := y - x_0$ at point x_0, by definition of hyperplane, $-a$ is a feasible direction as well.
Consequently, $a \cdot \nabla f(x_0) = 0$ for every feasible direction. That is, $\nabla f(x_0) \perp \Omega$. ■

Geometric Interpretation Let $d_0, d_1, \ldots, d_{n-1}$ be a set of non-zero Q-orthogonal vectors in $\mathbb{R}^n$. Let $B_k = Span\{d_0, \ldots, d_{k-1}\}$ for $k = 0, 1, \ldots, n$.
Note:

- $$B_0 = \{0\} \subseteq B_1 = \langle d_0 \rangle \subseteq B_2 = \langle d_0, d_1 \rangle \subseteq \ldots \subseteq B_n = \langle d_0, \ldots, d_{n-1} \rangle = \mathbb{R}^n$$
- $$\dim B_k = k$$
- $$x_0 + B_0 \subseteq x_0 + B_1 \subseteq \ldots$$

Theorem The sequence $\{x_k\}$ generated from $x_0 \in \mathbb{R}^n$ by conjugate directions method has the property that x_k minimizes $f(x) = \frac{1}{2}x^TQx - b^Tx$ on the affine hyperplane $x_0 + B_k$.

Proof. Recall that x_k is the minimizer of $f(x)$ on $x_0 + B_k \iff \nabla f(x_k) \perp x_0 + B_k$
Enough to prove that $\nabla f(x_k) \perp B_k$.

Remarks: x_0 here is like a bias which shifts the subspace by a "constant". Also, "$\nabla f(x_k) \perp B_k$" here means that $\nabla f(x_k)$ is perpendicular to every basis vector of B_k.

We prove this by induction on k.
Notation: $\nabla f(x_k) = Qx_k - b =: g_k$.
Base case: k = 0 $B_0 = \{0\} \implies g_0 \perp B_0$

Inductive Step: Assume that $g_k \perp B_k$, show $g_{k+1} \perp B_{k+1}$
Since

$$x_{k+1} = x_k + \alpha_k d_k$$

then

$$\underbrace{Q_{x_{k+1}} - b}_{g_{k+1}} = \underbrace{Q_{x_k} - b}_{g_k} + \alpha_k Q d_k$$

$$g_{k+1}^T B_k = \langle d_0, \ldots, d_{k-1} \rangle \tag{71}$$

$$g_{k+1}^T d_k = (\underbrace{g_k + \alpha_k Q d_k^T d_k}_{g_{k+1}})^T d_k \tag{72}$$

$$= g_k^T d_k + \alpha_k d_k^T Q d_k \tag{73}$$

$$= g_k^T d_k + (-\frac{g_k^T d_k}{d_k^T Q d_k}) d_k^T Q d_k \tag{74}$$

$$= 0 \tag{75}$$

This implies that $g_{k+1} \perp d_k$
For $0 \leq i < k$,

$$g_{k+1}^T \cdot d_i = (g_k + \alpha_k Q d_k)^T d_i \tag{76}$$

$$= \underbrace{g_k^T d_i}_{=0} + \underbrace{\alpha_k d_k^T Q d_i}_{=0} \tag{77}$$

$$= 0 \tag{78}$$

Therefore, $g_{k+1} \perp d_0, d_1, \ldots, d_k$
Hence $g_{k+1} \perp \langle d_0, d_1, \ldots, d_k \rangle = B_k$ ■

Corollary x_n minimizes $f(x)$ on $x_0 + B_n$ (which is $\mathbb{R}^n$)
i.e. $x_n = x^*$

Corollary $0 \leq q(x_k) = \min_{x \in x_0 + B_k} q(x) \leq q(x_{k-1}) = \min_{x \in x_0 + B_{k-1}} q(x)$

Corollary

$$\begin{cases} \min f(x) \\ x \in x_0 + B_1 \end{cases} \tag{79}$$

$$\implies \begin{cases} \min f(x_0 + t d_0) \\ t \in \mathbb{R} \end{cases} \qquad \text{(Since } x_0 + B_1 = \{x_0 + t d_0 | t \in \mathbb{R}\}\text{)}$$

$$\implies 0 = \frac{d}{dt}|_{t=t_0} f(x_0 + t d_0) = \nabla f(x_0 + t_0 d_0) \cdot d_0 \qquad \text{(where } t_0 \text{ is such that } x_1 = x_0 + t_0 d_0\text{)}$$

■

8 Calculus of Variations

Note: infinite dimensional optimization.

Comparison with finite dimensions

	finite dimensional	∞-dimensional
problem	$\min f(x)$	$\min F[u]$
constraint	$x \in M$	$u \in \mathcal{A}$
note	set of points in $\mathbb{R}^n$	space of functions

Model model

$$\mathcal{A} = \{u : [0,1] \to \mathbb{R} | u \in C^1 \, s.t. \, u(0) = u(1) = 1\}$$

Note: We call F a "Functional". It maps a function to a real number.

Notation Write $u(\cdot)$ for a function u.

8.1 Example

$F[u(\cdot)] = \frac{1}{2}\int_0^1 \{u(x)^2 + u'(x)^2\}\, dx$.

$\begin{cases} \min F[u(\cdot)] \\ u(\cdot) \in \mathcal{A} \end{cases}$

means: Find $u^*(\cdot) \in \mathcal{A}$ s.t. $F[u^*(\cdot)] \leq F[u(\cdot)]$ for all $u(\cdot) \in \mathcal{A}$.

Plan

1. We derive 1st order necessary conditions for a local min;
2. Find a function $u^*(\cdot)$ satisfying these conditions;
3. Check this candidate $u^*(\cdot)$ is in fact a minimizer.

We reduce this problem to (many) 1-dimensional problems.
Step 1: Derive 1st order necessary conditions for a local min
Fix $v(\cdot) \in C^1$ on $[0,1]$ s.t. $v(0) = 0 = v(1)$.
Suppose $u^*(\cdot) \in \mathcal{A}$ is a minimizer.
Notice that $u^*(\cdot) + sv(\cdot) \in \mathcal{A} \, \forall s \in \mathbb{R}$.
Let $f : \mathbb{R} \to \mathbb{R}$ s.t. $f(s) := F(u^*(\cdot) + sv(\cdot))$.
If $u^*(\cdot)$ minimizes F, then $s = 0$ minimizes f, then $f'(0) = 0$.
Then $f(0) = F[u^*(\cdot)] \leq F[u^*(\cdot) + sv(\cdot)] = f(s)$

$$f'(0) = \frac{d}{ds}|_{s=0} \underbrace{F[u^*(\cdot) + sv(\cdot)]}_{f(s)} \tag{80}$$

$$= \frac{d}{ds}|_{s=0} \frac{1}{2} \int_0^1 \{[u^*(x) + sv(x)]^2 + [u^{*'}(x) + sv'(x)]^2\}\, dx \tag{81}$$

$$= \frac{1}{2}\frac{d}{ds}|_{s=0} \int_0^1 \{u^*(x)^2 + u^{*'}(x)^2\}\, dx + \frac{d}{ds}|_{s=0} s \int_0^1 \{u^*(x)v(x) + u^{*'}(x)v'(x)\}\, dx + \frac{d}{ds}|_{s=0} \frac{s^2}{2} \int_0^1 \{v(x)^2 + v'(x)^2\}\, dx \tag{82}$$

$$= \int_0^1 \{u^*(x)v(x) + u^{*'}(x)v'(x)\}\, dx \tag{83}$$

So far, if $u^*(\cdot)$ is a minimizer of F over $\mathcal{A}$, then

$$\int_0^1 \{u^*(x)v(x) + u'^*(x)v'(x)\}\, dx = 0 \quad (\heartsuit)$$

for all $v(\cdot) \in C^1$ on $[0,1]$ and $v(0) = 0 = v(1)$. We call this a "primitive form of 1st order condition", and call $v(\cdot)$ the test functions.
Recall Integration by parts:

$$\int_0^1 w(x)v'(x)\,dx = w(x)v(x)|_0^1 - \int_0^1 w'(x)v(x)\,dx$$

$$(\heartsuit) = \int_0^1 u^*(x)v(x)\,dx + \int_0^1 u'^*(x)v'(x)\,dx \tag{84}$$

$$= \int_0^1 u^*(x)v(x)\,dx + \underbrace{u'^*(x)v(x)|_0^1}_{=0\;(v(0)=v(1)=0)} - \int_0^1 u''^*(x)v(x)\,dx \tag{85}$$

$$= \int_0^1 \big(u^*(x) - u''^*(x)\big)\,v(x)\,dx \tag{86}$$

$$= 0 \tag{87}$$

For all test functions $v(\cdot)$.

Lemma: Fundamental Lemma of Calculus of Variations Suppose g is continuous function on interval $[a,b]$. If

$$\int_a^b g(x)v(x)\,dx = 0$$

for all test functions $v(\cdot)$, then
$g(x) \equiv 0$ on $[a,b]$.
Then by Fundamental Lemma of Calculus of Variations, $(\heartsuit) \implies u^*(x) - u''^*(x) \equiv 0$, which is the 1st order necessary condition for $u^*(\cdot)$.
Step 2: Find a function $u^*(\cdot)$ satisfying these conditions

$$\begin{cases} u^*(x) = u''^*(x) \\ u^*(0) = u^*(1) = 1 \end{cases} \implies u^*(x) = c_1 e^x + c_2 e^{-x} \tag{88}$$

$$\begin{cases} 1 = u^*(0) = c_1 + c_2 \\ 1 = u^*(1) = c_1 e + c_2 \frac{1}{e} \end{cases} \implies c_1 = \frac{1}{e+1}, c_2 = \frac{e}{e+1} \tag{89}$$

$$\implies u^*(x) = \frac{1}{e+1}e^x + \frac{e}{e+1}e^{-x} \tag{90}$$

Step 3: check $u^*(\cdot)$ is in fact a global minimizer.
We derived that

$$F[u^*(\cdot) + sv(\cdot)] = F[u^*(\cdot)] + s\underbrace{\int_0^1 \{u^*(x)v(x) + u'^*(x)v'(x)\}\,dx}_{=0} + \underbrace{\frac{s^2}{2}\int_0^1 \{v(x)^2 + v'(x)^2\}\,dx}_{\geq 0}$$

$$F[u^*(\cdot)] \leq F[u^*(\cdot) + sv(\cdot)]$$

for all test functions $v(\cdot)$ and all $s \in \mathbb{R}$. In particular, let $s = 1$, then

$$F[u^*(\cdot)] \leq F[u^*(x) + v(\cdot)]$$

for all test functions $v(\cdot)$. In particular, let $v(\cdot) = u(\cdot) - u^*(\cdot)$, then

$$F[u^*(\cdot)] \leq F[u(\cdot)]$$

for all $u(\cdot) \in \mathcal{A}$.
Note: The space of $v(\cdot)$ is a vector space, but $\mathcal{A}$ is not a vector space (since $u(\cdot) \neq 0$). It is a translate of a vector space.

Lemma: Fundamental Lemma of Calculus of Variations Suppose g is continuous function on interval $[a, b]$. If

$$\int_a^b g(x)v(x)\,dx = 0$$

for all test functions $v(\cdot)$, then
$g(x) \equiv 0$ on $[a, b]$.

Proof. Suppose for contradiction that $g(x) \not\equiv 0$ on $[a, b]$.
WLOG, $g(x_0) > 0$ for some $x_0 \in (a, b)$. This implies that $g > 0$ on $(c, d) \subsetneq (a, b)$.
Let $v(\cdot)$ be a continuous function s.t.

$$v(\cdot)\begin{cases} > 0 & \text{on } (c, d) \\ = 0 & \text{otherwise} \end{cases}$$

Then

$$\int_a^b g(x)v(x)\,dx = \int_c^d \underbrace{g(x)v(x)}_{>0}\,dx > 0$$

which leads to a contradiction. ■

8.2 Classical Problem: the Brachistochrone

Galileo (1638): Find the curve connecting A and B on which a point mass moves without fiction under the influence of gravity in the least time possible.
Johann Bernoulli (1696): Revisit the problem and sent invitations
6 correct solutions sent (1697):
Leibniz, Johann, Jacob, l'Hospital, Von Tschinhaus, Anonymous → Newton (*)
This answer is the beginning of Calculus of Variations

8.3 General class of problems in Calculus of Variations

$$\mathcal{A} = \{u : [a, b] \to \mathbb{R} | u \in C^1, u(a) = A, u(b) = B\}$$

$$F[u(\cdot)] = \int_a^b L(x, u(x), u'(x))\,dx$$

where $L(x, z, p) : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$

Model example $L(x, z, p) = \frac{z^2+p^2}{2}$, $F[u(\cdot)] = \int_0^1 \frac{u(x)^2+u'(x)^2}{2}\,dx$
Notation:

$$\begin{cases} L_z(x, z, p) = \frac{\partial}{\partial z} L(x, z, p) \\ L_p(x, z, p) = \frac{\partial}{\partial p} L(x, z, p) \end{cases}$$

Definition Given $u(\cdot) \in \mathcal{A}$, suppose $\exists$ function $g(\cdot)$ on $[a, b]$ s.t.

$$\frac{d}{ds}|_{s=0} F[u(\cdot) + sv(\cdot)] = \int_a^b g(x)v(x)\,dx$$

for all test functions $v(\cdot)$, then $g(\cdot)$ is called the variational derivative of F at $u(\cdot)$,
denoted by $\frac{\delta F}{\delta u}(u)(\cdot)$ or $\frac{\delta F}{\delta u}(u)$ or $\frac{\delta F}{\delta u}$.

Analogy **In finite dimensions**:

$$\frac{d}{ds}|_{s=0} f(u+sv) = \nabla f(u) \cdot v \tag{91}$$

$$= \sum_i \nabla f(u)_i v_i \tag{92}$$

for all $v \in \mathbb{R}^n$

In Calculus of Variations (∞ dimensional):

$$\frac{d}{ds}|_{s=0} F[u(\cdot)+sv(\cdot)] = \int_a^b \frac{\delta F}{\delta u}(u)(x) v(x)\,dx \qquad \text{(where possible)}$$

$$\sim \sum_x \frac{\delta F(u)}{\delta u}(x) v(x) \qquad \text{(a kind of an infinite sum)}$$

Model example $L(x,z,p) = \frac{z^2+p^2}{2}$, $F[u(\cdot)] = \int_0^1 \frac{u(x)^2+u'(x)^2}{2}\,dx$

$$\frac{d}{ds}|_{s=0} F[u(\cdot)+sv(\cdot)] = \ldots = \int_0^1 [u(x)-u''(x)] v(x)\,dx$$

for all test functions $v(\cdot)$

Therefore $\frac{\delta F}{\delta u}(u)(x) = u(x) - u''(x)$

Lemma (1st order necessary conditions satisfied by a solution $u^*(\cdot) \in C^1$)

$$\mathcal{A} = \{u : [a,b] \to \mathbb{R} | u \in C^1, u(a) = A, u(b) = B\}$$

If $u^*(\cdot) \in \mathcal{A}$ minimizes F over $\mathcal{A}$, and if $\frac{\delta F}{\delta u}(u^*)(\cdot)$ exists and is continuous, then it must satisfy

$$\frac{\delta F}{\delta u}(u^*)(\cdot) \equiv 0$$

Proof. note: $u^*(\cdot) + sv(\cdot) \in \mathcal{A}$

If $u^*(\cdot)$ is a minimizer of F, then

$$F[u^*(\cdot) \le F[u^*(\cdot) + sv(\cdot)]$$

for all test functions v.

Define $f(s) := F[u^*(\cdot) + sv(\cdot)]$, then $f(0) \le f(s)$ for all $s \in \mathbb{R}$.

Then

$$\int_a^b \frac{\delta F}{\delta u}(u^*)(\cdot) v(x)\,dx = \frac{d}{ds}|_{s=0} F[u^*(\cdot) + sv(\cdot)] \qquad \left(\frac{\partial F}{\partial u}(u^*)(\cdot) \text{ exists}\right)$$

$$= \frac{d}{ds}|_{s=0} f(s) \tag{93}$$

$$= f'(0) = 0 \qquad \text{(0 is the global minimize of } f\text{)}$$

This implies that $\frac{\delta F}{\delta u}(u^*)(\cdot) \equiv 0$. ■

Theorem: Leibniz Integral Rule Let $f(x,t)$ be a function such that both $f(x,t)$ and its partial derivative $\frac{\partial}{\partial x} f(x,t)$ is continuous w.r.t. t and x in some region of the (x,t)-plane, including $a(x) \le t \le b(x), x_0 \le x \le x_1$. Also suppose that the functions $a(x)$ and $b(x)$ are both continuous and both have continuous derivatives for $x_0 \le x \le x_1$. Then, for $x_0 \le x \le x_1$,

$$\frac{d}{dx}\left(\int_{a(x)}^{b(x)} f(x,t)\,dt\right) = f(x,b(x)) \cdot \frac{d}{dx} b(x) - f(x,a(x)) \cdot \frac{d}{dx} a(x) + \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} f(x,t)\,dt$$

Note that if $a(x)$ and $b(x)$ are constants rather than functions of x, we have a special case of Leibniz's rule:

$$\frac{d}{dx}\left(\int_a^b f(x,t)\,dt\right) = \int_a^b \frac{\partial}{\partial x} f(x,t)\,dt$$

Theorem

$$\mathcal{A} = \{u : [a,b] \to \mathbb{R} | u \in C^1, u(a) = A, u(b) = B\}$$

$$F[u(\cdot)] = \int_a^b L(x, u(x), u'(x))\, dx$$

where $L \in C^2$.
Then if $u(\cdot) \in C^1$, then $\frac{\delta F}{\delta u}(u)(\cdot)$ exists, is continuous, and

$$\frac{\delta F}{\delta u}(u)(x) = -\frac{d}{dx}[L_p(x, u(x), u'(x))] + L_z(x, u(x), u'(x))$$

Let v be a test function ($v(a) = v(b) = 0$)

$$\frac{d}{ds}|_{s=0} F[u(\cdot) + sv(\cdot)] = \frac{d}{ds}|_{s=0} \int_a^b L(x, u(x) + sv(x), u'(x) + sv'(x))\, dx \tag{94}$$

$$= \int_a^b \frac{d}{ds}|_{s=0} L(x, z, p)\, dx \quad \text{(By Leibniz's rule)}$$

$$= \int_a^b L_z(\cdot)v(x) + L_p(\cdot)v'(x)\, dx \tag{95}$$

$$= \int_a^b L_z(\cdot)v(x)\, dx + \int_a^b L_p(\cdot)v'(x)\, dx \tag{96}$$

$$= \int_a^b L_z(\cdot)v(x)\, dx + L_p(\cdot)v(x)|_a^b - \int_a^b \frac{d}{dx} L_p(\cdot)v(x)\, dx \quad \text{(Integration by parts)}$$

$$= \int_a^b [-\frac{d}{dx} L_p(\cdot) + L_z(\cdot)]v(x)\, dx \ \forall \text{ test functions } v(\cdot) \tag{97}$$

By the definition of variational derivative, it follows that

$$\frac{\delta F}{\delta u}(u)(x) = -\frac{d}{dx}[L_p(x, u(x), u'(x))] + L_z(x, u(x), u'(x))$$

Furthermore, since $L(\cdot) \in C^2, -\frac{d}{dx}L_p(\cdot)$ and $L_z(\cdot)$ are continuous. Moreover, $u(\cdot)$ and $u'(\cdot)$ are continuous, so is the composite function. Hence the variational derivative is continuous.

Model example $L(x, z, p) = \frac{z^2+p^2}{2}, \quad F[u(\cdot)] = \int_0^1 \frac{u(x)^2+u'(x)^2}{2}\, dx$
$L_z(x, z, p) = z \implies L_z(x, u(x), u'(x)) = u(x)$
$L_p(x, z, p) = p \implies L_p(x, u(x), u'(x)) = u'(x)$

$$\frac{\delta F}{\delta u}(u(\cdot)) = -\frac{d}{dx}[u'(x)] + u(x) = -u''(x) + u(x)$$

If $u^*(\cdot) \in \mathcal{A}$ is a minimizer, then $-u''(x) + u(x) = 0$

Example 8.1 (min arclength)**.** We will show that the straight line gives the shortest path.

$$\min\ F[u(\cdot)] = \int_a^b (1 + u'(x)^2)^{\frac{1}{2}}\, dx = \text{arclength of } u(\cdot)$$

$$\mathcal{A} = \{u : [a,b] \to \mathbb{R} | u \in C^1, u(a) = A, u(b) = B\}$$

Then $L(x, z, p) = (1 + p^2)^{\frac{1}{2}}, L_z = 0$ and $L_p = \frac{p}{(1+p^2)^{\frac{1}{2}}}$
If $u^*(\cdot)$ is a minimizer, then

$$-\frac{d}{dx}\frac{u'(x)}{(1 + u'(x))^{\frac{1}{2}}} \equiv 0$$

$\implies \frac{u'(x)}{(1+u'(x))^{\frac{1}{2}}} = const$
$\implies u'(x)^2 = const(1 + u'(x)^2)$
Then $u'(x) = \alpha$ for some $\alpha \in \mathbb{R}$.
Then $u(x) = \alpha x + \beta$ for some $\beta \in \mathbb{R}$.

Example 8.2 (Surface Area of Revolution)**.** Suppose $u(\cdot) \in C^1$ on $[a, b]$, the surface area of rotating the curve u connecting a and b can be computed as

$$F[u(\cdot)] = \int_a^b 2\pi u(x)\sqrt{1 + u'(x)}\, dx$$

For simplicity, assume $u > 0$. In this example, the space of feasible functions is

$$\mathcal{A} = \{u : [a, b] \to \mathbb{R} : u \in C^1, u(a) = A, u(b) = B, u > 0\}$$

If $u(\cdot)$ solves the minimization problem, it must be the case that

$$\frac{\delta F}{\delta u}(u)(\cdot) \equiv 0 \quad (\dagger)$$

Notice

$$L(x, z, p) = 2\pi z\sqrt{1 + p^2}$$
$$L_z(x, z, p) = 2\pi\sqrt{1 + p^2}$$
$$L_p(x, z, p) = 2\pi z \frac{p}{\sqrt{1 + p^2}}$$

Claim: the family of $u(\cdot) = \beta \cosh(\frac{x-\alpha}{\beta})$ solves the necessary condition †.
Instance 1 When $a = 0, b = 1, A = B = 1$, plugging in the initial condition gives

$$\begin{cases} \beta \cosh\left(\frac{0-\alpha}{\beta}\right) & = 1 \\ \beta \cosh\left(\frac{1-\alpha}{\beta}\right) & = 1 \end{cases} \tag{98}$$

solving above system of equations provides the solution.
Instance 2 When $a = 0, b = 1, A = 1, B = 0$, plugging in these initial conditions gives

$$\begin{cases} \beta \cosh\left(\frac{0-\alpha}{\beta}\right) & = 1 \\ \beta \cosh\left(\frac{1-\alpha}{\beta}\right) & = 0 \end{cases} \tag{99}$$

because $\cosh > 0$, the second equation suggests $\beta = 0$, but in this case the first equation would never hold. Therefore, there is no solution to this calculus of variation.
In face, the surface area is minimized by

$$u(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases} \tag{100}$$

8.4 Euler-Lagrange Equations in $\mathbb{R}^n$

Setup

$$F[u(\cdot)] = \int_a^b L(x, u(x), u'(x))\, dx \tag{101}$$
$$u : [a, b] \to \mathbb{R}^n \tag{102}$$
$$L(x, z, p) : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \tag{103}$$
$$\mathcal{A} := \left\{u : [a, b] \to \mathbb{R}^n : u \in C^1, u(a) = \mathbf{A}, u(b) = \mathbf{B}\right\} \tag{104}$$

Theorem 8.1 (Euler-Lagrange Equations in Vector Forms)**.**

$$-\frac{d}{dx}\nabla_p L(x,z,p) + \nabla_z L(x,z,p) = \mathbf{0} \in \mathbb{R}^n \quad (\dagger) \tag{105}$$

Example 8.3 (Classical Lagrangian Mechanics)**.**

$$V(x) : \mathbb{R}^n \to \mathbb{R} \text{ potential energy} \tag{106}$$

$$\frac{1}{2}m||v||_2^2 \text{ kinetic energy} \tag{107}$$

$$L(t,x,v) := \frac{1}{2}m||v||_2^2 - V(x) \text{ difference between KE and PE} \tag{108}$$

Consider a path $x(t)$ in $\mathbb{R}^n$, define objective function as

$$F[x(\cdot)] = \int_a^b L(t, x(t), x'(t))\ dt \tag{109}$$

$$= \int_a^b \frac{1}{2}m||\dot{x}(t)||_2^2 - V(x(t))\ dt \tag{110}$$

The Euler-Lagrange equation in vector form implies

$$-\frac{d}{dt}\nabla_{(3)}L(t, x(t), \dot{x}(t)) + \nabla_{(2)}L(t, x(t), \dot{x}(t)) = 0 \tag{111}$$

$$\implies -\frac{d}{dt}m\dot{x}(t) - \nabla V(x(t)) = 0 \tag{112}$$

$$\implies m\ddot{x}(t) = \nabla V(x(t)) \quad (\dagger\dagger) \tag{113}$$

Remark 8.1. ($\dagger\dagger$) is often referred to as *Newton's second law*: object moves along the path on which the total conversion between kinetic and potential energies is minimized.

Example 8.4 (3-Dimensional Pendulum)**.** Suppose the pendulum is moving on a path such that the total conversion between kinetic and potential energies is minimized, that is

$$\min \int_a^b L(t)\ dt = \int_a^b \frac{1}{2}m(\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2) - mgz(t)\ dt \tag{114}$$

with the restriction that $||\mathbf{x}(t)|| = \ell$, where ℓ is the radius of the sphere.
The restriction can be embodied by framing the problem using spherical coordinates:

$$x := \ell \cos\varphi \sin\theta \tag{115}$$

$$y := \ell \sin\varphi \sin\theta \tag{116}$$

$$z := -\ell \cos\theta \tag{117}$$

where the path of motion can be characterized using $(\theta(t), \varphi(t))$.
The objective function is therefore

$$L\left(t, \begin{pmatrix}\theta(t)\\ \varphi(t)\end{pmatrix}, \begin{pmatrix}\dot{\theta}(t)\\ \dot{\varphi}(t)\end{pmatrix}\right) = \frac{1}{2}m\ell^2(\dot{\theta}^2 + \dot{\varphi}^2\sin^2(\theta)) + mg\ell\cos\theta \tag{118}$$

So the Euler-Lagrange equation can be written as

$$-\frac{d}{dt}\nabla_{(3)}L\left(t, \begin{pmatrix}\theta(t)\\ \varphi(t)\end{pmatrix}, \begin{pmatrix}\dot{\theta}(t)\\ \dot{\varphi}(t)\end{pmatrix}\right) + \nabla_{(2)}L\left(t, \begin{pmatrix}\theta(t)\\ \varphi(t)\end{pmatrix}, \begin{pmatrix}\dot{\theta}(t)\\ \dot{\varphi}(t)\end{pmatrix}\right) = \mathbf{0} \tag{119}$$

8.5 Equality constraints

8.5.1 Isoperimetric constraints

Recall: finite dimensional case

$$f, g : \mathbb{R}^n \to \mathbb{R} \tag{120}$$

$$\begin{cases} \min\limits_{x \in \mathbb{R}^n} & f(x) \\ & g(x) = const \end{cases} \tag{121}$$

Suppose regular point x_* ($\nabla g(x_*) \neq 0$) is a minimizer. Then $\exists \lambda \in \mathbb{R}$ s.t. $\nabla f(x_*) + \lambda \nabla g(x_*) = 0$ (by Lagrange multipliers)

Remark 8.2. x_* minimizes $f + \lambda g$. The Lagrange multipliers convert the original problem to an unconstrained optimization problem $L(x, \lambda) = f(x) + \lambda g(x)$.

Infinite dimensional case

$$F[u(\cdot)] = \int_a^b L^F(x, u(x), u'(x))\, dx \tag{122}$$

$$G[u(\cdot)] = \int_a^b L^G(x, u(x), u'(x))\, dx \tag{123}$$

$$\tag{124}$$

$$\begin{cases} \min\limits_{u(\cdot) \in \mathcal{A}} & F[u(\cdot)] \\ & G[u(\cdot)] = const \end{cases}$$

Suppose regular point $u_*(\cdot)$ ($\frac{\delta G}{\delta u}(u_*) \neq 0$) is a minimizer, then $\exists \lambda \in \mathbb{R}$ s.t.

$$\frac{\delta F}{\delta u}(u_*) + \lambda \frac{\delta G}{\delta u}(u_*) \equiv 0$$

Remark 8.3. u_* minimizes $F + \lambda G$.

Example 8.5.

$$\mathcal{A} := \{u : [-a, a] \to \mathbb{R}, u \in C^1, u(-a) = u(a) = 0\} \tag{125}$$

$$F[u(\cdot)] = \int_a^b u(x)\, dx \tag{126}$$

$$G[u(\cdot)] = \int_a^b \sqrt{1 + u'(x)}\, dx = l > 0 \quad \text{note that } G \text{ is arg length} \tag{127}$$

$$\begin{cases} \min\limits_{u \in \mathcal{A}} & (-F)[u(\cdot)] \\ & G[u(\cdot)] = l \end{cases} \tag{128}$$

Let $u_*(\cdot)$ be a minimizer, then

$$\frac{\delta F}{\delta u} = -\frac{d}{dx} L_p^F + L_z^F$$
$$\frac{\delta G}{\delta u} = -\frac{d}{dx} L_p^G + L_z^G$$

Then Euler-Lagrange equations suggests

$$-\frac{d}{dx} L_p^F + L_z^F + \lambda(-\frac{d}{dx} L_p^G + L_z^G) = 0$$

$$\implies \lambda^2 \frac{u_*'(x)^2}{1 + u_*'(x)^2} = (C_1 - x)^2 (\dagger)$$

Claim: Solution $u_*(\cdot)$ to (†) satisfies

$$(x - C_1)^2 + (u_*(x) - C_2)^2 = \lambda^2$$

So that the graph of $u_*(\cdot)$ lies on a circle, and our solution is the semi-circle that has length l.
Check:

$$\frac{d}{dx}\left[2(x - C_1) + 2(u_*(x) - C_2)u_*'(x)\right] = 0 \tag{129}$$

which implies

$$u_*'(x) = -\frac{x - C_1}{u_*(x) - C_2} \tag{130}$$

$$\implies u_*'(x)^2 = \frac{(x - C_1)^2}{(u_*(x) - c_2)^2} \quad (\S) \tag{131}$$

Also,

$$(u_*'(x)^2)(u_*(x) - C_2)^2 = (x - C_1)^2 + (u_*(x) - C_2)^2 = \lambda^2 \tag{132}$$

$$\implies (u_*(x) - C_2)^2 = \frac{\lambda^2}{1 + u_*'(x)^2} \quad (\S\S) \tag{133}$$

Combine (§) and (§§),

$$\frac{\lambda^2}{1 + u_*'(x)^2} u_*'(x)^2 = (x - C_1)^2 \tag{134}$$

It is possible to solve for u:

$$\begin{cases} x = -a, y = 0, (-a - C_1)^2 + (0 - C_2)^2 = \lambda^2 \\ x = +a, y = 0, (+a - C)1)^2 + (0 - C_2)^2 = \lambda^2 \\ \int_{-a}^{a} \sqrt{1 + u'(x)^2}\, dx = l \end{cases}$$

8.5.2 Holonomic constraints

Setup(3-Dim Special Case) Minimize

$$F[x(\cdot), y(\cdot), z(\cdot)] = \int_a^b L(t, x(t), y(t), z(t), \dot{x}(t), \dot{y}(t), \dot{z}(t))\, dt$$

with constraint

$$H(x(t), y(t), z(t)) \equiv 0$$

Theorem 8.2 (Euler-Lagrange Equations)**.** Let $\mathbf{x}_*(t) := \begin{pmatrix} x_*(t) \\ y_*(t) \\ z_*(t) \end{pmatrix}$ be the minimizer subject to the constraint, then

$$\begin{pmatrix} \frac{\delta F}{\delta x}[x_*(\cdot), y_*(\cdot), z_*(\cdot)](t) \\ \frac{\delta F}{\delta y}[x_*(\cdot), y_*(\cdot), z_*(\cdot)](t) \\ \frac{\delta F}{\delta z}[x_*(\cdot), y_*(\cdot), z_*(\cdot)](t) \end{pmatrix} + \lambda(t) \begin{pmatrix} H_x[x_*(\cdot), y_*(\cdot), z_*(\cdot)](t) \\ H_y[x_*(\cdot), y_*(\cdot), z_*(\cdot)](t) \\ H_z[x_*(\cdot), y_*(\cdot), z_*(\cdot)](t) \end{pmatrix} = 0 \quad \forall t \in \mathbb{R}$$

where $\lambda : [a, b] \to \mathbb{R}$ is a function.