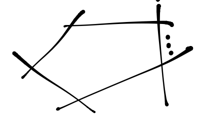
Ch 11. Constrained optimization

$\min\limits_{x \in \mathbb{R}^n} f(x)$ s.t. $h_1(x)=0$, $h_2(x)=0$, $\vdots$, $h_m(x)=0$; $g_1(x)\le 0$, $\vdots$, $g_p(x)\le 0$

$m \le n$

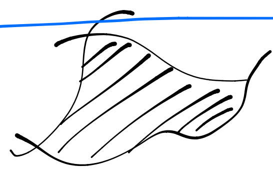
if doesn't matter for whether $p>n$ or $p<n$ or $p=n$

can even be infinitely many



Example:

$\min\limits_{x,y\in\mathbb{R}} f(x,y)$ s.t. $x+y\ge 0$, $x+4y-4\le 0$, $y\le 4$



inequality constrain

equality constain (if > dim, might have no solution)

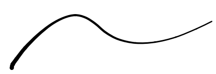
Def: An active constrain is a constrain that is equality at optimal point. Otherwise it is called inactive.

$\begin{pmatrix} h_1(x) \\ \vdots \\ h_m(x) \end{pmatrix}$ $h:\mathbb{R}^n\to\mathbb{R}^n$, $h(x)=0$ $g:\mathbb{R}^n\to\mathbb{R}^n$, $g(x)\le 0$

Example

$h_1(x)=\cdots=h_m(x)=0$

$h(x)=|x|-1$ is a sphere



Definition:

if $h:\mathbb{R}^n\to\mathbb{R}^m$, if h is "regular" then $h(x)=0$ is an hyper-surface of dim $n-m$ $(h:\mathbb{R}^n\to\mathbb{R}^m)$ (regular = matrix with rank n s.t. dim ker $=n-m$)

Definition:

$S=\{x: h(x)=0\}$

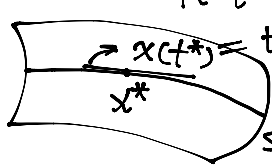
A curve on S is a family of point $x(t)\in S$ continuously parametrized by $t\in[a,b]$

$\frac{d}{dt}x(t)=\dot{x}(t)$

Example:

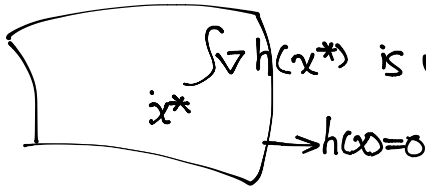
$x(t^*)=x^*$

$\dot{x}(t^*)=\frac{d}{dt}\Big|_{t=t^*}x(t)$ is a tangent vector of S at x^*



Def: Consider all differential curve $x(t)\in S$ s.t. $x(t^*)=x^*$ for a t^*. The tangent plane at x^* will be the set of derivative at x^*.

Example: $M=\{y:\nabla h(x^*)y=0\}$



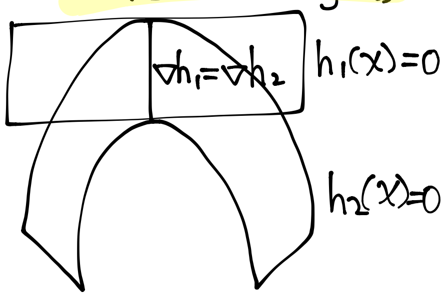
$\nabla h(x^*)$ is orthogonal to the surface at x^*

Def: A point x^* s.t. $h(x^*)=0$ is a regular point of $\{h(x)=0\}$ if $\nabla h_1(x^*), \cdots, \nabla h_m(x^*)$ are lin. independent.

Note: if h is linear then $h(x)=Ax+b$, $\nabla h = A$, regular $\Leftrightarrow A$ has full rank (rank n)

A is m×n matrix

Thm: At a regular point, x^* of $\{h=0\}$, the tangent plane is equal to $M=\{y: \nabla h(x^*)y=0\}$



Tangent plane $\supset M$

Proof: T = tangent plane, we want $T=M$
will prove $T \subset M$ and $M \subset T$

$T \subset M$

Let $y \in T$

$y = x'(t^*)$ for $x(t) \in S$

$h(x(t))=0$

$\frac{d}{dt}(h(x(t)') = 0$

$\nabla h(x(t)) \cdot x'(t) = 0$

$\nabla h(x(t))y = 0 \Rightarrow y \in M \Rightarrow T \subset M$

$M \subset T$

Let $y \in M$, $\nabla h(x^*)y = 0$

we want $x(t)$ s.t. $x(t^*) = x^*$ and $x'(t^*) = y$

suppose $t^*=0$, $x(t) = x^* + ty + \nabla h(x^*)u(t)$

we want to know if $\exists u$ s.t.

$h(x^* + ty + \nabla h(x^*)^T u(t)) = 0$

$\nabla_u h|_{t=0} \neq 0$

choose $u(0)=0$, $\nabla h(x^*) \cdot \nabla h(x^*)^T \neq 0$

since full rank, by implicit function theorem $\exists u$ near $u=0$

s.t. $h(x^* + ty + \nabla h(x^*)u) = 0$

$\frac{d}{dt}\Big|_{t=0} h(x(t)) = 0$

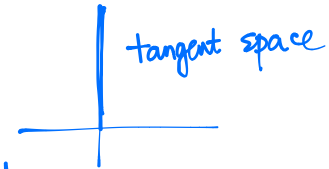
$\nabla h(x^*)y + \underbrace{\nabla h(x^*)\nabla h(x^*)^T}_{\neq 0} h'(0) = 0$

$\Rightarrow h'(0) = 0$

$$x'(0) = y + \nabla h(x^*)^T h'(0) = y$$
$$\Rightarrow y \in T \Rightarrow M \subset T$$

Example: $h(x,y) = x$ $\quad\quad$ $\bar{h}(x,y) = x^2$

$\nabla h = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\quad\quad$ $\nabla \bar{h} = \begin{pmatrix} 2x \\ 0 \end{pmatrix}$



$M_h = \left\{ \begin{pmatrix} u \\ v \end{pmatrix} : \begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} = 0 \right\} = \left\{ \begin{pmatrix} 0 \\ v \end{pmatrix} \right\}$ on the tangent plane,

$M_{\bar{h}} = \left\{ \begin{pmatrix} u \\ v \end{pmatrix} : \begin{pmatrix} 0 \\ 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} = 0 \right\} = \mathbb{R}^2$ not regular not on the tangent plane

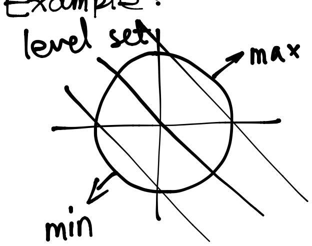
1st order necessary condition

recall: $\min_{x \in \Omega} f(x)$ $\quad$ 1st order NC: $\nabla f(x^*) d \geq 0 \ \forall$ feasiable d,

d is feasible $\Leftrightarrow \begin{cases} \nabla h(x^*) d = 0 \\ \nabla h(x^*)(-d) = 0 \end{cases}$

$\Rightarrow \nabla f(x^*) d = 0$ since $\begin{cases} \nabla f(x^*) d \geq 0 \\ \nabla f(x^*)(-d) \geq 0 \end{cases}$

Example:



$\min \ x + y = f(x,y)$ s.t. $x^2 + y^2 = 1$

$h(x,y) = x^2 + y^2 - 1$

$\nabla h = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$

Lemma: let x^* be a regular point of $h(x) = 0$ and a local min of f s.t. $h(x) = 0$. Then $\forall y \in \mathbb{R}^n$ s.t. $\nabla h(x^*) y = 0$, we have $\nabla f(x^*) y = 0$

Proof: Let y s.t. $\nabla h(x^*) y = 0$

y is in the tangent space

$\exists \ x(t) \in \{h(x) = 0\}$

s.t. $x(0) = x^*$ $\quad$ $x'(0) = y$

x^* is a min $\Rightarrow 0$ is a min

$$\frac{d}{dt}\Big|_{t=0} f(x(t)) = 0 \Leftrightarrow \nabla f(x(0))\, x'(0) = 0$$
$$\Leftrightarrow \nabla f(x^*) y = 0$$

Theorem: Let x^* be a local min of f s.t. $\underbrace{h(x) = 0}_{M \text{ equations}}$ if x^* is regular for the constraint, then

$\exists \lambda \in \mathbb{R}^m$ s.t. $\nabla f(x^*) + \lambda^T \nabla h(x^*) = 0$

we call λ the Lagrange Multiplier.

$\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ $\ (f : \mathbb{R}^n \to \mathbb{R}) \in \mathbb{R}^n$

$\lambda^T \nabla h(x) = 1 \times m \quad m \times n = 1 \times n \quad \in \mathbb{R}^n$

$\Rightarrow \nabla f(x^*) + \lambda^T \nabla h(x^*) = 0$ is a system of n equations

$\Rightarrow$ total $m+n$ equations with $n(x) + m(\lambda)$ variables

Definition : The Lagrangian associated with a min problem

$l(x,\lambda) = f(x) + \lambda^T(x)$

$\left.\begin{array}{l} \nabla_x l = \nabla f(x) + \lambda^T \nabla h(x) = 0 \\ \nabla_\lambda l = h(x) \qquad\qquad\quad = 0 \end{array}\right\} \Rightarrow$ 1 st order NC

$\nabla l = 0 \Rightarrow \nabla_x l = \nabla_\lambda l = 0$

Example p328

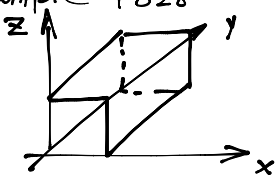
min $x_1x_2 + x_2x_3 + x_1x_3$ s.t. $x_1+x_2+x_3=3$

$l(x_1,x_2,x_3,\lambda) = x_1x_2+x_2x_3+x_1x_3 + \lambda(x_1+x_2+x_3-3)$

$$\left.\begin{array}{l} \nabla_x l = \begin{pmatrix} x_2+x_3+\lambda \\ x_3+x_1+\lambda \\ x_1+x_2+\lambda \end{pmatrix} = 0 \\ \text{with } x_1+x_2+x_3-3=0 \end{array}\right\} \begin{array}{l} x_1=x_2=x_3=1 \\ \lambda=-2 \end{array}$$

of "λ" depends on the # of "s.t" in this case, one λ.

Example P328



max xyz s.t. $2xy+2yz+2zx=c$

$x \geq 0$

$y \geq 0$

$z \geq 0$

$l(x,y,z,\lambda) = xyz + \lambda(2xy+2zx+2yz-c)$

$$\nabla l = \begin{pmatrix} yz+2y\lambda+2z\lambda \\ xz+2\lambda(x+z) \\ xy+2\lambda(x+y) \\ 2(xy+xz+yz)-c \end{pmatrix} \begin{array}{c} ① \\ ② \\ ③ \\ ④ \end{array} = 0$$

- ① + ② + ③ $= xy+xz+yz+4\lambda(x+y+z) = \frac{c}{2} + 4\lambda(x+y+z) = 0$

$\Rightarrow \lambda \neq 0$.

- Suppose $x=0$, ② gives us $2\lambda z = 0$, so $z=0$

③ gives us $2\lambda y = 0$, so $y=0$

But $2(xy+yz+zx) = c$, so $x \neq 0$

- Similarly if we suppose $y=0$ and $z=0$, we will have the same contradiction $\Rightarrow$ $x\neq0, y\neq0, z\neq0$.

- $x① - y② = 0 \Rightarrow 2\lambda(x-y)=0$, $x=y$

Similarly, we can solve to get $x=y=z$

- substitute back, $6x^2=c$, $x=\sqrt{\frac{c}{6}}$

Example : min x_2+x_3 s.t. $\begin{cases} x_1+x_2+x_3=1 \\ x_1^2+x_2^2+x_3^2=1 \end{cases}$

Let $y=x_2=x_3$ since x_2 and x_3 have the same ???(missing a word)

$\Rightarrow$ min $2y$ s.t. $\begin{cases} x+2y=1 \\ x^2+2y^2=1 \end{cases}$

$l(x,y,\lambda_1,\lambda_2) = 2y + \lambda_1(x+2y-1) + \lambda_2(x^2+2y^2-1)$

$$\nabla l = \begin{pmatrix} \lambda_1 + 2\lambda_2 x \\ 2 + 2\lambda_1 + 4\lambda_2 y \\ x + 2y - 1 \\ x^2 + 2y^2 - 1 \end{pmatrix}$$

$\begin{cases} y = 0 \\ x = 1 \end{cases}$ or $\begin{cases} y = \frac{2}{3} \\ x = \frac{-1}{3} \end{cases}$

$1, 0, 0$

$-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}$

original $$\nabla l = \begin{pmatrix} \lambda_1 + 2\lambda_2 x_1 \\ 1 + \lambda_1 + 2\lambda_2 x_2 \\ 1 + \lambda_1 + 2\lambda_2 x_3 \\ x_1 + x_2 + x_3 - 1 \\ x_1^2 + x_2^2 + x_3^2 - 1 \end{pmatrix}$$

$$\nabla l(1, 0, 0, \lambda_1, \lambda_2) = \begin{pmatrix} \lambda_1 + 2\lambda_2 \\ 1 + \lambda_1 \\ 1 + \lambda_1 \\ 0 \\ 0 \end{pmatrix}$$

$\exists \lambda_1 = -1, \lambda_2 = -\frac{1}{2}$

so $1, 0, 0$ works

$\nabla l(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}, \lambda_1, \lambda_2)$ ... works as well