Linear Constraints

$$\text{minimize } f(x) \quad \text{s.t. } Ax = b \qquad \text{— (P)}$$

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

$$h_i(x) = a_i^T x - b_i, \quad i = 1, \dots, m$$

$$\nabla h_i(x) = a_i$$

If a_i's are l.i. (for which m must be $\leq n$)
Then all $x \in \mathbb{R}^n$ are regular, and if x^* is a local min, then corresponding LM λ satisfies:

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x^*) = 0$$

i.e. $$\sum_{i=1}^{m} \lambda_i a_i = -\nabla f(x^*)$$

i.e. $$A^T \lambda = -\nabla f(x^*)$$

Also, x^* satisfies:

$$Ax^* = b$$

If $m = n$ then A is an $n \times n$ matrix, and x^* is unique:

$$x^* = A^{-1} b$$

and this is the solution to (P) since this is the only feasible choice for x.

If $m < n$, A is an $m \times n$ matrix with rank m, then we can have infinitely many solutions to $Ax^* = b$.
Need to solve for x^* and λ together using

$$\left.\begin{aligned} A^T \lambda + \nabla f(x^*) &= 0 \\ Ax^* &= b \end{aligned}\right\} \begin{array}{l} m+n \text{ equations} \\ m+n \text{ unknowns } \lambda, x^*. \end{array}$$

If $\nabla f(x^*)$ is linear in x^*, then solution can be unique. In general, we may have multiple solutions for x^* and corresponding λ.

Check second order sufficiency condition:

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i (a_i^T x - b_i)$$

$$\nabla_x L(x, \lambda) = \nabla f(x) + \sum_{i=1}^{m} \lambda_i a_i$$

$$\nabla_{xx}^2 L(x, \lambda) = \nabla^2 f(x)$$

Need to check:

$$z^T \nabla^2 f(x^*) z > 0$$

for all $z \neq 0$ s.t.

$$a_i^T z = 0, \quad i = 1, \dots, m.$$

Example minimize $-(x_1 x_2 + x_2 x_3 + x_1 x_3)$
s.t.

$$\left.\begin{aligned} x_1 + x_2 &= 2 \\ x_2 + x_3 &= 1 \end{aligned}\right\} \quad Ax = b, \; A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \; b = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

$$f(x) = -(x_1 x_2 + x_2 x_3 + x_1 x_3)$$

$$h_1(x) = x_1 + x_2 - 2$$

$$h_2(x) = x_2 + x_3 - 1$$

$$\nabla f(x) = -\begin{bmatrix} x_2 + x_3 \\ x_1 + x_3 \\ x_1 + x_2 \end{bmatrix}, \quad \nabla h_1(x) = \overset{a_1}{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}, \quad \nabla h_2(x) = \overset{a_2}{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}}$$

x^*, λ Satisfy : $A x^* = b, \quad A^T \lambda + \nabla f(x^*) = 0$

$$\left.\begin{aligned} x_1^* + x_2^* &= 2 \\ x_2^* + x_3^* &= 1 \\ \lambda_1 - (x_2^* + x_3^*) &= 0 \\ \lambda_1 + \lambda_2 - (x_1^* + x_3^*) &= 0 \\ \lambda_2 - (x_1^* + x_2^*) &= 0 \end{aligned}\right\} \Rightarrow \begin{aligned} \lambda_1 &= 1 \\ \lambda_2 &= 2 \\ x_1^* + x_3^* &= 3 \end{aligned}$$

$$2(x_1^* + x_2^* + x_3^*) = 3 + 2 + 1 = 6$$

$$\Rightarrow x_1^* + x_2^* + x_3^* = 3$$

$$\Rightarrow x_3^* = 1, \quad x_1^* = 2, \quad x_2^* = 0$$

Check second order Suff. Condition for $x^* = (2, 0, 1)$:

$$\nabla^2_{xx} L(x, \lambda) = \nabla^2 f(x)$$

$$= \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$$

Consider principal minors : $|0| = 0 \quad \begin{vmatrix} 0 & -1 \\ -1 & 0 \end{vmatrix} = -1 < 0$

$\Rightarrow \nabla^2_{xx} L(x^*, \lambda)$ not PD or even PSD.

But, we only need $z^T \nabla^2_{xx} L(x^*, \lambda) z > 0$
for $z \neq 0$ s.t. $a_i^T z = 0$, $i = 1, 2$.

$$a_1^T z = [1 \ 1 \ 0] z = z_1 + z_2 = 0 \Rightarrow z_1 = -z_2$$
$$a_2^T z = [0 \ 1 \ 1] z = z_2 + z_3 = 0 \Rightarrow z_3 = -z_2$$

$$z^T \nabla^2_{xx} L(x^*, \lambda) z = [z_1 \ z_2 \ z_3] \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$
$$= [z_1 \ z_2 \ z_3] \begin{bmatrix} -z_2 - z_3 \\ -z_1 - z_3 \\ -z_2 - z_3 \end{bmatrix}$$
$$= -2 (z_1 z_2 + z_2 z_3 + z_1 z_3)$$
$$= 2 z_2^2 > 0 \quad \text{if } z \neq 0$$

Thus $x^* = (2, 0, 1)$ is a (strict) local min of $f(x)$ s.t. $h_i(x) = 0$, $i = 1, 2$.

But is $x^* = (2, 0, 1)$ a global min?
Does f have a global min under $h(x) = 0$?

$h(x) = 0 \Leftrightarrow x_1 = 2 - x_2 \quad x_3 = 1 - x_2$

$$f(x) = -(x_1 x_2 + x_2 x_3 + x_1 x_3)$$
on $h(x)=0$
$$= -((2 - x_2) x_2 + (1 - x_2) x_2 + (1 - x_2)(2 - x_2))$$
$$= -(2 - x_2 + x_2 - x_2^2) = x_2^2 - 2$$

$\Rightarrow$ f is coercive on $\mathcal{H} = \{x : h(x) = 0\}$ which is closed
$\Rightarrow$ by Corollary to WT, f achieves global min on $\mathcal{H}$

$\Rightarrow$ $x^* = (2, 0, 1)$ is the global min of $f(x)$ on $\mathcal{H}$.

Linear Constraints with no regular x

$$\begin{aligned} &\text{minimize } f(x) \\ &\text{s.t. } Ax = b \end{aligned} \qquad A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \quad m \le n \qquad a_i^T x = b_i, \; i = 1, \dots, m$$

Suppose $a_1, a_2, \dots, a_m$ are not l.i.

Then $\nabla h_i(x) = a_i \Rightarrow \{\nabla h_i(x)\}_{i=1}^m$ not l.i.

$\Rightarrow$ no regular x.

$a_1, a_2, \dots, a_m$ not l.i. $\Rightarrow$ $\operatorname{rank}(A) = k < m$.

Reorder the a_i's so that $a_1, a_2, \dots, a_k$ are l.i.

Then

$$a_i = \sum_{j=1}^{k} \alpha_j^{(i)} a_j, \quad i = k+1, \dots, m$$

Then for $Ax = b$ to have at least one solution (consistency), we must have

$$b_i = \sum_{j=1}^{k} \alpha_j^{(i)} b_j, \quad i = k+1, \dots, m$$

Thus the constraints $a_i^T x = b_i$, $i = k+1, \dots, m$ are redundant and can be discarded.

$$\begin{aligned} &\text{minimize } f(x) \\ &\text{s.t. } Ax = b \end{aligned} \quad \equiv \quad \begin{aligned} &\text{minimize } f(x) \\ &\text{s.t. } \tilde{A} x = \tilde{b} \end{aligned}$$

with $\tilde{A} = \begin{bmatrix} a_1^T \\ \vdots \\ a_k^T \end{bmatrix}$, $\tilde{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_k \end{bmatrix}$

$\Rightarrow$ LM's exist

But $\tilde{A}, \tilde{b}$ may not be unique $\Rightarrow$ λ may not be unique for a given x^*

<u>Example</u> minimize $x_1^2 + x_2^2 + 2x_3^2$

s.t. $x_1 + x_2 = 2$

$-x_2 + x_3 = 1$

$x_1 + x_3 = 3$

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \rightarrow \text{rows are l.d.}$$

$$\nabla f(x) = [2x_1 \quad 2x_2 \quad 4x_3]^T$$

<u>Choice 1</u> $\tilde{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$ $\quad \bar{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

x^*, $\lambda = (\lambda_1, \lambda_2)$ satisfy:

$\tilde{A}^T \lambda = -\nabla f(x^*)$, i.e.,

$$\left.\begin{aligned} \lambda_1 &= -2x_1^* \\ \lambda_1 - \lambda_2 &= -2x_2^* \\ \lambda_2 &= -4x_3^* \end{aligned}\right\} \Rightarrow -2x_2^* = 4x_3^* - 2x_1^*$$

$$\Rightarrow x_2^* = x_1^* - 2x_3^*$$

and

$x_1^* + x_2^* = 2$

$x_3^* - x_2^* = 1$

$$\left.\begin{aligned} x_1^* + x_2^* &= 2 \\ x_1^* - x_2^* - 2x_3^* &= 0 \end{aligned}\right) \Rightarrow 2x_2^* + 2x_3^* = 2$$

$$\Rightarrow x_2^* = 1 - x_3^*$$

$x_3^* - (1 - x_3^*) = 1 \Rightarrow 2x_3^* = 2 \Rightarrow x_3^* = 1$

$\Rightarrow x_2^* = 0$

$\Rightarrow x_1^* = 2$

$\lambda_1 = -4$, $\lambda_2 = -4$.

Choice 2 $\quad \tilde{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \tilde{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

x^* and λ satisfy :

$$\left.\begin{aligned} \lambda_1 + \lambda_2 &= -2x_1^* \\ \lambda_1 &= -2x_2^* \\ \lambda_2 &= -4x_3^* \end{aligned}\right\} \Rightarrow x_1^* = x_2^* + 2x_3^*$$

$$\begin{aligned} x_1^* + x_2^* &= 2 \\ x_1^* + x_3^* &= 3 \end{aligned}$$

$$\left.\begin{aligned} x_1^* + x_2^* &= 2 \\ x_1^* - x_2^* - 2x_3^* &= 0 \end{aligned}\right\} \Rightarrow 2x_1^* - 2x_3^* = 2$$

$$\Rightarrow x_1^* - x_3^* = 1$$

$$\left.\begin{aligned} x_1^* + x_3^* &= 3 \\ x_1^* - x_3^* &= 1 \end{aligned}\right\} \Rightarrow \begin{aligned} x_1^* &= 2 \\ x_3^* &= 1 \end{aligned}$$

$$\Rightarrow x_2^* = 0$$

So x^* is same as in choice 1,

but $\lambda_1 = 0$, $\lambda_2 = -4$