Proposition 3.1.1 (Lagrange Multiplier Theorem – Necessary Conditions) NEP

Let x^* be a local minimum of 'f' subject to $h(x)=0$, and assume that the constraint gradients $\nabla h_1(x^*) \cdots \nabla h_m(x^*)$ are linearly independent. Then there exists a unique vector $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)$ called Lagrange multiplier vector, such that

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

$$\boxed{\begin{array}{ll} & \min f(x) \\ \text{s.t.} & h_1(x)=0 \\ & h_2(x)=0 \\ & h_m(x)=0 \end{array}}$$

If in addition f and h are twice continuously differentiable, we have

$L(x,\lambda) = f(x) + \lambda_1 h_1(x) + \lambda_2 h_2(x) \cdots + \lambda_m h_m(x)$

$$y^T \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x^*) \right) y \geq 0 \quad \forall\, y \in V(x^*)$$

where $V(x^*)$ is the subspace of the first order variations

$$V(x^*) = \{ y \mid \nabla_i h(x^*)^T y = 0, \ i=1 \cdots m$$

$Z^T \nabla^2 f(x^*) Z$ is s.p.d. (semidefinite)

$$A = \begin{bmatrix} \nabla h_1(x^*)^T \\ \nabla h_m^T(x^*) \end{bmatrix} \quad \underline{\underline{Z}}$$

$\underset{m \times n}{A} x = b$

$a_1^T x = b_1$
$a_2^T x = b_2$
$\vdots$
$a_m^T x = b_m$

these are linearly independent

$\underset{n \times m}{A^T}$ matrix has rank m.

all the columns of A^T are l.i.

$h_1(x) = a_1^T x - b_1 = 0$
$h_2(x) = a_2^T x - b_2 = 0$
$\vdots$
$h_m(x) = a_m^T x - b_m = 0$

$$\left.\begin{array}{l} \nabla h_1(x) = a_1 \\ \nabla h_2(x) = a_2 \\ \vdots \\ \nabla h_m(x) = a_m \end{array}\right\} \text{regularity condition.}$$

$$\underset{n \times m}{\left[\nabla h_1(x), \nabla h_2(x) \cdots \nabla h_m(x) \right]}$$

rank of above matrix is not m; then Lagrange multiplier theorem is not applicable and we need to check optimality from a bottoms up approach)

$$\min \; f(x)$$
$$h(x) = 0 \qquad m\text{-eqns}$$

$$\mathcal{L}(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x)$$ is a function in n+m variables

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0;$$
$$\nabla_\lambda \mathcal{L}(x^*, \lambda^*) = 0;$$
} optimal is the stationary point of the Lagrangian function.

$$\frac{\partial \mathcal{L}}{\partial x_1} = 0 \Rightarrow \left.\frac{\partial f}{\partial x_1}\right|_{x^*} + \sum_{i=1}^{m} \lambda_i^* \frac{\partial h_i}{\partial x_1} = 0$$
$$\vdots$$
$$\frac{\partial \mathcal{L}}{\partial x_n} = 0 \Rightarrow \left.\frac{\partial f}{\partial x_n}\right|_{x^*} + \sum_{i=1}^{m} \lambda_i^* \frac{\partial h_i}{\partial x_n} = 0$$

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

$$\frac{\partial \mathcal{L}}{\partial \lambda_1} = 0 \Rightarrow h_1(x^*) = 0;$$
$$\frac{\partial \mathcal{L}}{\partial \lambda_2} = 0 \Rightarrow h_2(x^*) = 0;$$
$$\vdots$$
$$\frac{\partial \mathcal{L}}{\partial \lambda_m} = 0 \Rightarrow h_m(x^*) = 0$$
} x^* the optimal should be a feasible point.

$$\mathcal{L}(x, \lambda^*) = f(x) + \lambda_1^* h_1(x) + \lambda_2^* h_2(x) + \cdots + \lambda_m^* h_m(x)$$

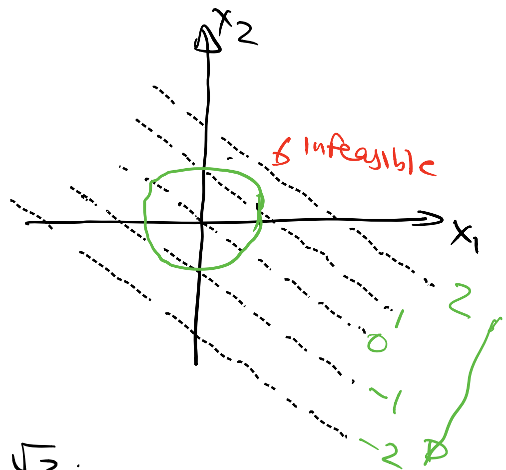
$$\frac{\partial^2 \mathcal{L}}{\partial x_i \partial x_j} = \nabla^2 f(x) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x)$$
$i = 1 \ldots m$

$$\nabla^2 f(x) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right]_{n \times n} \qquad \nabla h_i(x) = \left[\frac{\partial^2 h_i(x)}{\partial x_i \partial x_j}\right]_{n \times n}$$

curvature information

$\min\ x_1 + x_2$

sub $x_1^2 + x_2^2 = 2$

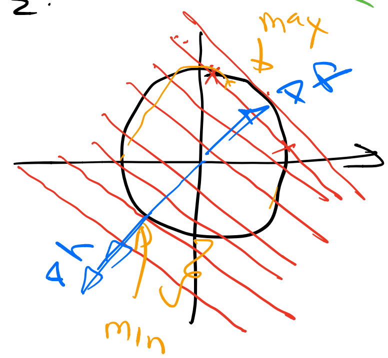


$x_1 + x_2 = 0$

$x_1 = -x_2$

$x_1^2 + x_2^2 = 2$

circle $(0,0)$, and radius $\sqrt{2}$.



$x^* = (-1, -1)$

$$\nabla f(x^*) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\nabla h(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}$$

$$\nabla h(x^*) = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$$

$$\nabla f(x^*) + \lambda \nabla h(x^*) = 0$$

$$\lambda = +\frac{1}{2}$$

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} -2 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Ex 3.1.1 [Bertsekas] [To illustrate, problem if regularity condition is not satisfied]

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

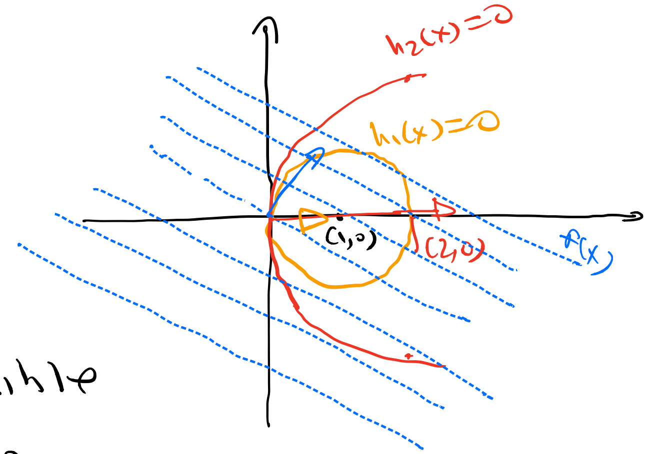
s.t.

$$h_1(x) = (x_1-1)^2 + x_2^2 - 1 = 0$$

$$h_2(x) = (x_1-2)^2 + x_2^2 - 4 = 0$$

$$(x_1-1)^2 + (x_2-0)^2 = 1^2$$

$$(x_1-2)^2 + (x_2-0)^2 = 2^2$$



Q] Does this problem has a optimal pt?

$$x^* = (0,0)$$

There is only feasible point $(0,0)$

$$\nabla h_1(x^*) = 2\begin{bmatrix} x_1^*-1 \\ x_2^* \end{bmatrix} = 2\begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

$\nabla h_1(x^*) \ldots \nabla h_m(x^*)$

$$\nabla h_2(x^*) = 2\begin{bmatrix} x_1^*-2 \\ x_2^* \end{bmatrix} = 4\begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Q] $\nabla h_1(x^*)$ and $\nabla h_2(x^*)$ are l.i.?

A] No!

$$2\nabla h_1(x^*) - \nabla h_2(x^*) = 0;$$