Constrained Optimization and Lagrange Multiplier

Zengchang Qin (PhD)

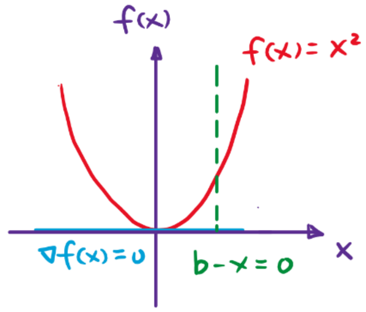
CONSTRAINED OPTIMIZATION

1) Find minimum or maximum of a function subject to some given constraints.

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

$$\text{s.t. } h_i(x)=0 \;, \; i=1,2\cdots\ell$$

$$g_j(x)\le 0 \;, \; j=1,2\cdots m$$



2) For unconstrained optimization.

$$f : \Omega \to R$$

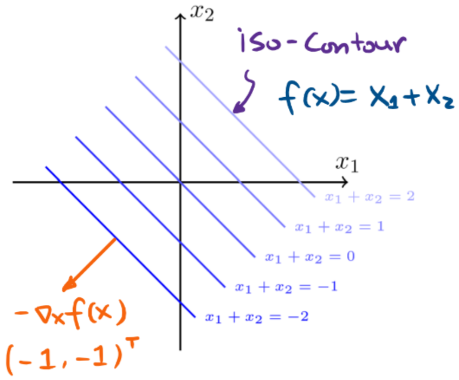
We aim to find $\nabla_x f(x)=0$

For example : $f(x)=x^2$

It is easy to see that the minimum is at $\nabla f(x)=0 \Rightarrow 2x=0 \Rightarrow x^*=0$

If we are given a constraint $h(x)$ where $h(x)=b-x=0$

Then, the function $x=b$ is called the feasible region.

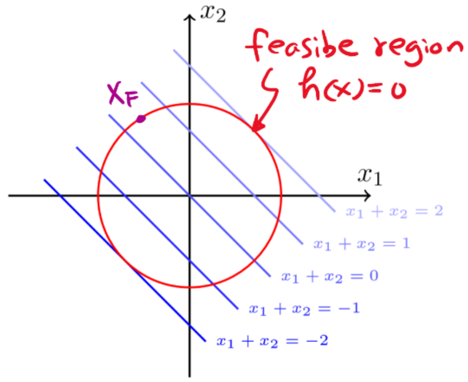


3) Consider a function $f(x)=x_1+x_2$

The iso-contour is drawn.

$$\nabla_x f(x) = \begin{bmatrix} \nabla_{x_1} \frac{\partial f}{\partial x_1} \\ \nabla_{x_2} \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

We cares about $-\nabla_x f(x) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

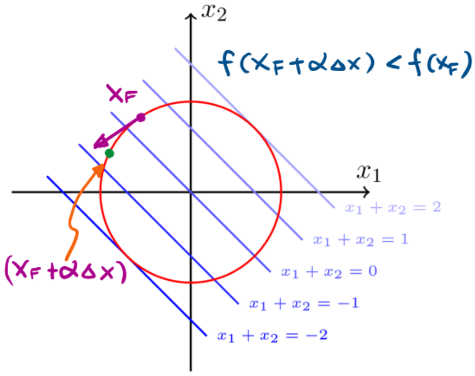


which is the direction of finding the minimum. (Gradient towards to the minimum)

4) Given the Constraint:

$$h(x) = x_1^2 + x_2^2 - 2$$

It is drawn in red (circle), the feasible region is the red circle.



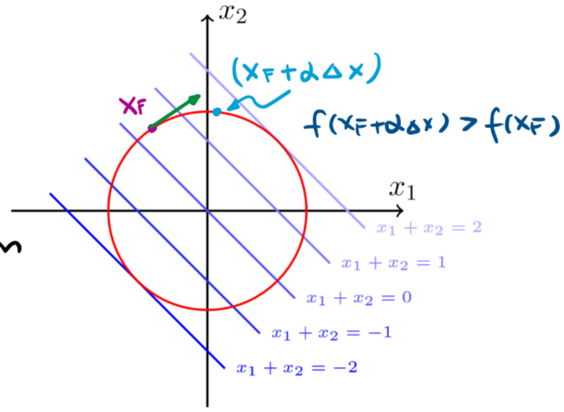
Given a point x_F on from the feasible region, we can move Δx so that

$$h(x_F + \alpha \Delta x) = 0$$

(move on the circle) and

$$f(x_F + \alpha \Delta x) < f(x_F)$$

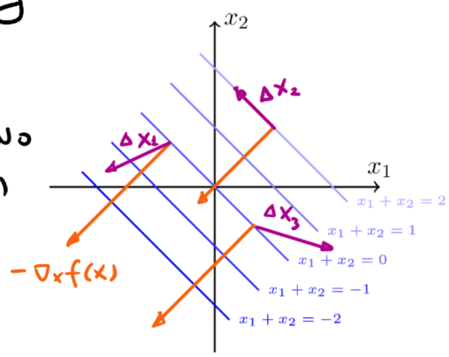
How can we find the correct direction to move?



5) In order to move x_F to a lower gradient we need to satisfy the following:

$$\Delta x\,(-\nabla_x f(x_F)) > 0$$

Consider the dot multiplication of two vectors. For the three cases given in the figure.



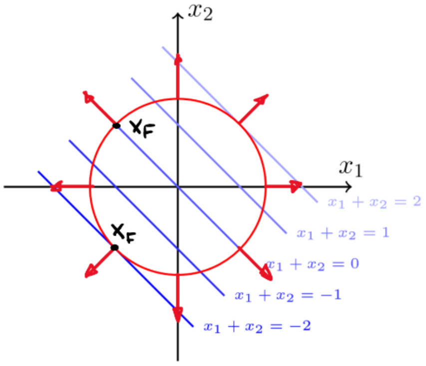
$$\begin{cases} -\nabla_x f(x) \cdot \Delta x_1 > 0 \\ -\nabla_x f(x) \cdot \Delta x_2 = 0 \\ -\nabla_x f(x) \cdot \Delta x_3 < 0 \end{cases}$$

6) $h(x) = x_1^2 + x_2^2 - 2$

$$\nabla_x h(x) = \begin{bmatrix} \frac{\partial h}{\partial x_1} \\ \frac{\partial h}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}$$

E.g. $x_F = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \Rightarrow \nabla_x h(x) = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$

$x_F = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \Rightarrow \nabla_x h(x) = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$



For any given point x_F in the feasible region

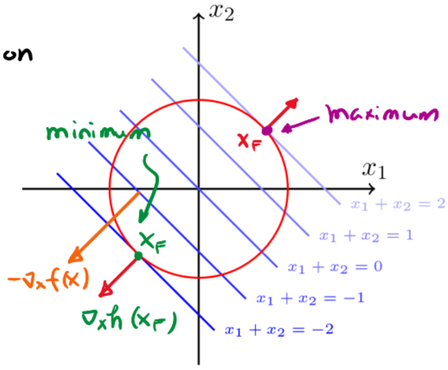
$$\begin{cases} \Delta x \cdot (-\nabla_x f(x_F)) > 0 \\ x_F \leftarrow x_F + \alpha \Delta x \end{cases}$$

Till $\Delta x \cdot (-\nabla_x f(x_F)) = 0$ that means

$\Delta x \perp \nabla_x h(x_F) \Rightarrow \Delta x \cdot \nabla_x h(x_F) = 0$

$$-\nabla_x f(x_F) = \mu \nabla_x h(x_F)$$

(Parallel direction)



7) Minimize $f(x)$ subject to $h(x) = 0$, we need to find x^*

$$-\nabla_x f(x^*) = \mu \nabla_x h(x^*) \qquad \text{Eq (1)}$$

We can construct Lagrange multiplier μ

$$L(x, \mu) = f(x) + \mu h(x)$$

$$\begin{cases} \frac{\partial L}{\partial x} = \nabla_x f(x) + \mu \nabla_x h(x) = 0 \\ \frac{\partial L}{\partial \mu} = h(x) = 0 \end{cases}$$

For more than one equality constraints

$$L(x, \mu) = f(x) + \sum_{i=1}^{\ell} \mu_i h_i(x)$$