Lecture 20 [Lagrange Method]

Monday, 14 March 2022 1:38 PM

$$\begin{cases} \min \ f(\underline{x}) \\ \text{s.t.} \quad \underline{h}(\underline{x}) = \underline{0} \\ \underline{h}(\underline{x}) = \begin{bmatrix} h_1(\underline{x}) \\ h_2(\underline{x}) \\ \vdots \\ h_m(\underline{x}) \end{bmatrix} \end{cases}$$

$x \in \mathbb{R}^n$

$f : \mathbb{R}^n \to \mathbb{R}$

$\underline{h} : \mathbb{R}^n \to \mathbb{R}^m \quad (m \le n)$

$\underline{h}, \underline{f} \in \mathcal{C}^{(1)}$

Defn: [Regular point] A point $\underline{x}^*$ satisfying $h_1(\underline{x}^*) = h_2(\underline{x}^*) = \dots h_m(\underline{x}^*) = 0$ is said to be a regular point of the constraints, if the gradient vectors $\nabla h_1(\underline{x}^*), \nabla h_2(x^*), \dots, \nabla h_m(\underline{x}^*)$ are linearly independent.

Jacobian matrix

$$D\underline{h}(\underline{x}^*) = \begin{bmatrix} Dh_1(\underline{x}^*) \\ \vdots \\ Dh_m(\underline{x}^*) \end{bmatrix} = \begin{bmatrix} (\nabla h_1(\underline{x}^*))^T \\ \cdots \\ (\nabla h_m(\underline{x}^*))^T \end{bmatrix}$$

$\underline{x}^*$ is regular $\iff$ rank $(D\underline{h}(\underline{x}^*)) = m \quad [m \le n]$.

Ex.

$$\begin{bmatrix} h_1(\underline{x}) = x_1 \\ h_2(\underline{x}) = x_2^2 - x_3^2 \end{bmatrix}$$

$$\nabla h_1(\underline{x}) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\nabla h_2(x) = \begin{bmatrix} 0 \\ 2x_2 \\ -2x_3 \end{bmatrix}$$

$$D\underline{h}(\underline{x}) = \begin{bmatrix} 1 & 0 \\ 0 & 2x_2 \\ 0 & -2x_3 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2x_2 & -2x_3 \end{bmatrix}$$

$$Dh(x) = \begin{bmatrix} 0 & 2x_2 \\ 0 & -2x_3 \end{bmatrix} \qquad \begin{bmatrix} 0 & 2x_2 & -2x_3 \end{bmatrix}$$

$x^* = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ is a regular pt because (i) $h_1(x^*) = 0$, $h_2(x^*) = 0$.

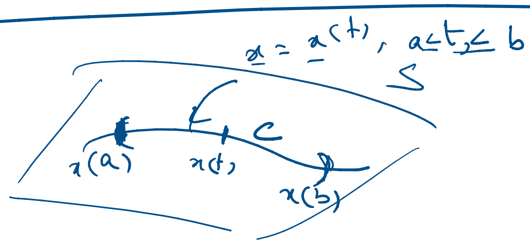
(ii) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -2 \end{bmatrix}$ l.i

$x^* = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ not a regular point because (ii) is not satisfied.

The set $h_1(x) = h_2(x) = \dots = h_m(x) = 0$ $\quad h_i : \mathbb{R}^n \to \mathbb{R}$
describes the surface

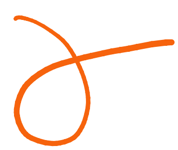
$$S = \{ x \in \mathbb{R}^n : h_1(x) = \dots = h_m(x) = 0 \}.$$

$\boxed{\text{Exercise}}$ If the points are regular, then $\dim(S) = n - m$.



Def [Curve]

A curve on S is a set of points $\{ x(t) \in S : t \in (a,b) \}$ continuously parametrized by $t \in (a,b)$; that $x : (a,b) \to S$ is a continuous function.



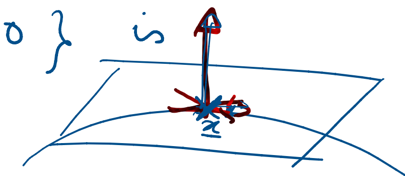
The curve $C = \{ x(t) : t \in (a,b) \}$ is differentiable if

$$\dot{x}(t) = \frac{dx}{dt} = \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \vdots \\ \dot{x}_n(t) \end{bmatrix}$$ exist for all $t \in (a,b)$.

The curve $C = \{ x(t) : t \in (a,b) \}$ is twice differentiable, The curve $C := \{ \underline{x}(t) : t \in (a, b) \}$ is <u>twice differentiable</u>,

if $\ddot{\underline{x}}(t) = \dfrac{d^2 \underline{x}}{dt^2} = \begin{bmatrix} \ddot{x}_1(t) \\ \vdots \\ \ddot{x}_n(t) \end{bmatrix}$ exists for all $t \in (a, b)$.

Defn. [<u>Tangent space</u>] at a point $\underline{x}^*$ on the surface $S := \{ \underline{x} \in \mathbb{R}^n : \underline{h}(\underline{x}) = 0 \}$ is



$$\boxed{T(\underline{x}^*) = \{ \underline{y} : \overset{\text{normal}}{D\underline{h}}(\underline{x}^*) \cdot \underline{y} = 0 \}}$$

$$= \mathcal{N}(D h(\underline{x}^*)) \subset \mathbb{R}^n.$$

Example: Let $S = \{ \underline{x} \in \mathbb{R}^3, \ h_1(\underline{x}) = x_1 = 0, \ h_2(\underline{x}) = x_1 - x_2 = 0 \}$

Find the tangent space of S.

$$D\underline{h}(\underline{x}) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}$$

$$D h(\underline{x}) \cdot \underline{y} = 0 \quad \Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$R_2 \to R_2 - R_1 \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$0 = y_1 \to$ basic

$0 = y_2 \to$ basic.

$y_3 \to$ free variable

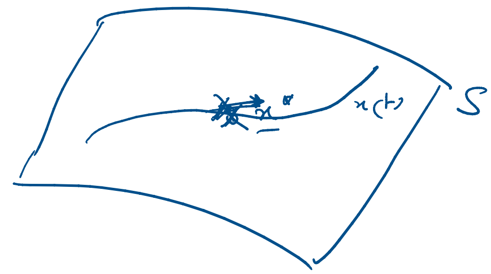
(Gauss-elimination)

$$\underline{y} = \left\{ \begin{bmatrix} 0 \\ 0 \\ \alpha \end{bmatrix} : \alpha \in \mathbb{R} \right\}$$

$\to \boxed{\text{Z-axis}}$

Thm [20.1 in Chong | Zak]
(Without proof).

Let $\underline{x}^* \in S$ be a regular point and $T(\underline{x}^*)$ be the tangent space at $\underline{x}^*$. Then $\underline{y} \in T(\underline{x}^*) \Leftrightarrow$ $\exists$ a differentiable curve in $\underline{S}$ passing through $\underline{x}^*$ with derivative $\underline{y}$ at $\underline{x}^*$.



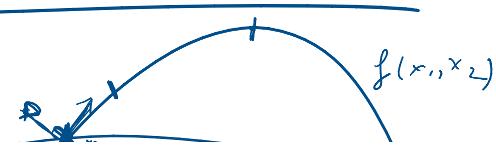
Defn. [Normal space] $N(\underline{x}^*)$ at a point $\underline{x}^*$ on the surface $S = \{\underline{x} \in \mathbb{R}^n : h(\underline{x}) = 0\}$ is the set

$$N(\underline{x}^*) = \{\underline{x} \in \mathbb{R}^n : \underline{x} = \underbrace{Dh(\underline{x}^*)^T}_{n \times m} \underbrace{\underline{z}}_{m \times 1}, \ \underline{z} \in \mathbb{R}^m\}$$

$$N(\underline{x}^*) = \mathcal{R}\left(D\underline{h}(\underline{x}^*)^T\right)$$
$$= \text{span}\{\nabla h_1(\underline{x}^*), \nabla h_2(\underline{x}^*), \ldots \nabla h_m(\underline{x}^*)\}$$
$$= \{\underline{x} \in \mathbb{R}^n : \underline{x} = z_1 \nabla h_1(\underline{x}^*) + \cdots + z_m \nabla h_m(\underline{x}^*), \ z_1, \ldots, z_m \in \mathbb{R}\}$$

$$\mathbb{R}^n = N(\underline{x}^*) \oplus T(\underline{x}^*)$$
$$T(\underline{x}^*) = N(\underline{x}^*)^{\perp}$$
$$N(\underline{x}^*) = T(\underline{x}^*)^{\perp}$$

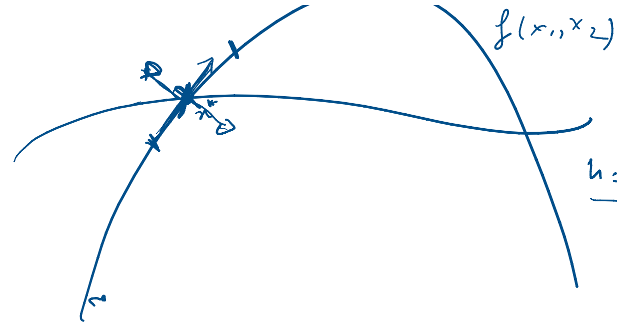
(Exercise) to revise.



... condition

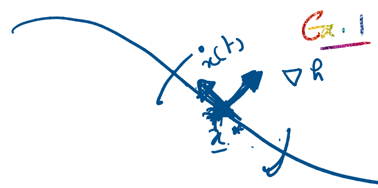
Lagrange condition

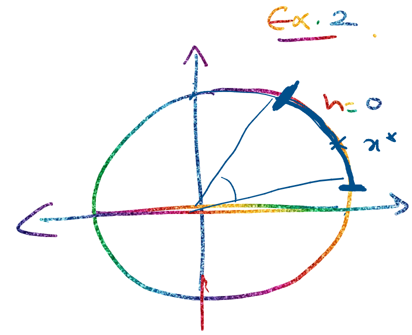
$\left[\begin{array}{l} f \quad \text{fn of two variables} \\ h : \mathbb{R}^2 \to \mathbb{R} \quad \to \text{one constraint.} \end{array}\right.$



$h = 0$

$x_1^2 + x_2^2 = 1$





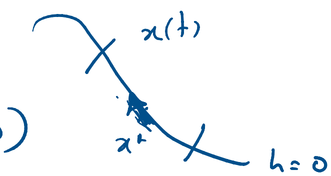
Parametrize the level set in the neighbourhood of x^* by $\{\underline{x}(t)\}$ that is a continuous differentiable function $\underline{x} : \mathbb{R}^2 \to \mathbb{R}$

s.t. $\underline{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$, $t \in (a, b)$

$\underline{x}^* = x(t^*)$, $\dot{x}(t^*) \neq 0$, $t^* \in (a, b)$.

Implicit fn theorem.

$$\boxed{\nabla h(\underline{x}^*) \perp \dot{\underline{x}}(t^*)}$$ Claim.



$h = 0 \Rightarrow h\left(x(t)\right) = 0 \quad t \in (a, b)$

$\nabla h \cdot \dot{x}(t) = 0$

$$\boxed{\nabla h(\underline{x}^*) \perp \dot{\underline{x}}(t^*)} \to$$

Let $\underline{x}^*$ be a minimizer of $f : \mathbb{R}^2 \to \mathbb{R}$ on the set $\{\underline{x} : h(\underline{x}) = 0\}$.

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

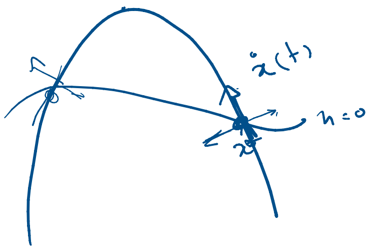
$\nabla h(x^*)$ is orthogonal to $\dot{\underline{x}}(t^*)$

the Set $\{ \underline{x} : h(\underline{x}) \ldots \}$

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

Claim: $\boxed{\nabla f(x^*) \text{ is orthogonal to } \dot{\underline{x}}(t^*)}$

$\phi(t) = f(x(t))$ attains a minimum at t^*.



$\phi'(t^*) = 0$

$$\boxed{\nabla f(x^*) \cdot \dot{x}(t^*) = 0}$$

$$\nabla f(x^*) \perp \dot{x}(t^*)$$

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

$\boxed{\text{Lagrange } \times}$

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

$$\ell(\underline{x}, \lambda) = f(\underline{x}) + \lambda h(\underline{x})$$

$\ell_{\underline{x}}(\underline{x}^*, \lambda^*) = 0$

$\ell_\lambda(\underline{x}^*, \lambda^*) = 0 \longrightarrow h(\underline{x}^*) = 0.$

f

$F(x^*) \succ 0$

Recall Hessian

$F > 0$ [minimizer]

$F < 0$ [maximizer]

$F = 0$] $\longrightarrow$ Inconclusive.

$$\underline{y}^T \mathcal{L}(\underline{x}^*) \underline{y}$$

$T(x^*)$