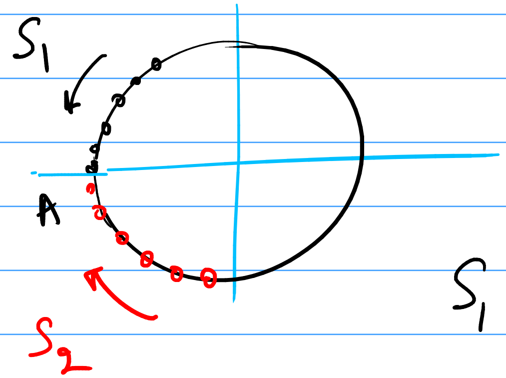
Context → Bring algebra & geometry into sync.

"Tangent Cone"

① Feasible sequence : Given an $x \in \Omega$

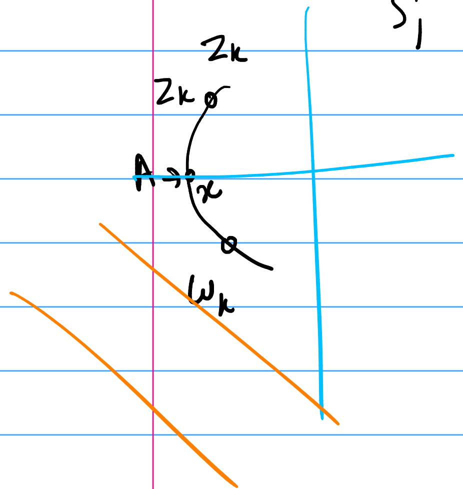
$\{Z_k\}$ is a F.S. if a) $Z_k \in \Omega$, b) $Z_k \to x$ } for Large k. $k \geq 1$.



$f(x) = x_1 + x_2$, $c_1(x) = x_1^2 + x_2^2 - 2 = 0$

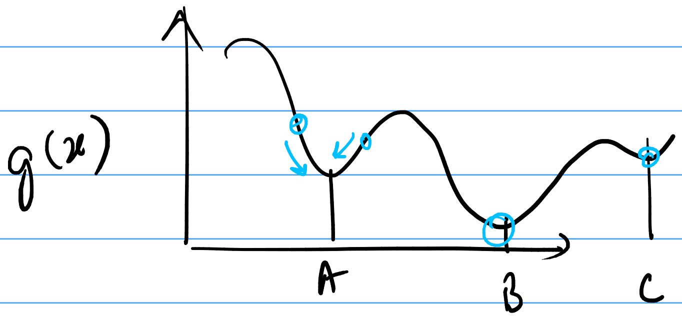
$$S_1 = \begin{bmatrix} -\sqrt{2 - 1/k^2} \\ 1/k \end{bmatrix}, \quad S_2 = \begin{bmatrix} -\sqrt{2 - 1/k^2} \\ -1/k \end{bmatrix}$$

② Local minimizer → A point x at which all feasible sequences satisfy $f(x) \leq f(z_k)$ → Large k.

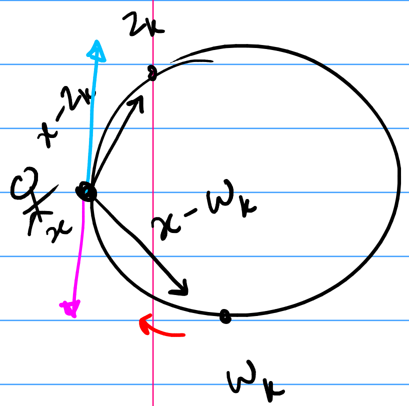


$S_1 \to \{z_k\}$, $S_2 \to \{w_k\}$, $f(x) = x_1 + x_2$

$$\left.\begin{array}{l} f(x) \leq f(z_k) \\ f(x) \geq f(w_k) \end{array}\right\} \Rightarrow$$ A is not a local minimizer.



③ Tangent → A limiting direction of a F.S.



In addition to $\{z_k\}$, I define a +ve scalar

d $\{t_k\}$ → has a property [$\lim_{k \to \infty} t_k = 0$]

$$d = \lim_{k \to \infty} \frac{x - z_k}{t_k}, \quad t_k = \|z_k - x\|_2$$

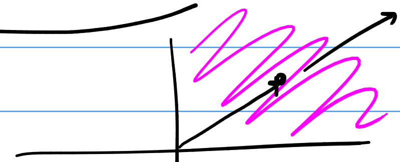
↳ Tangent Cone: Set of all tangents to Ω at a feasible point x, is called the T.C. $T_\Omega(x)$. At A, we have $T_\Omega(x) = \begin{bmatrix} 0 \\ d_2 \end{bmatrix}$, $d_2 \in \mathbb{R}$.

Aside

Cone:

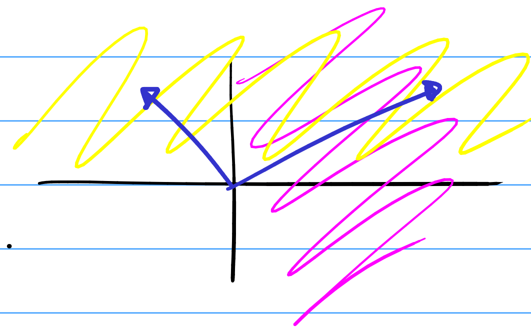
A set M s.t. $\forall\, x \in M$, $\alpha x \in M$ $\forall\, \alpha > 0$

eg. 1) $\left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \middle| \; x_1 > 0,\; x_2 > 0 \right\}$



Convex Cone

eg 2) $\left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \middle| \; x_1 \geq 0 \text{ OR } x_2 \geq 0 \right\}$



nonconvex cone.

↳ Tangent cone?

Say $d \in T_\Omega(x)$, then $\boxed{\alpha d \overset{?}{\in} T_\Omega(x)}$

$$d = \lim_{k \to \infty} \frac{z_k - x}{t_k} \qquad t_k' = t_k/\alpha$$

$t_k > 0$ $\quad t_k \to 0$ for large k. $\quad \to \alpha d.$

↳ Defn: LICQ: Linear independence constraint qualification: defn

Given a feasible pt x & the active constraint set $A(x)$, we say that LICQ holds if the set of active constraint gradients $\{\nabla c_i(x) \mid i \in A(x)\}$ is linearly indepn.

↳ Reln between algebra & geometry

Linearized feasible directions — Tangent cone

$$\nabla c_1 = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$$

Following are true:

① $T_\Omega(x) \subset \mathcal{F}(x)$

② If LICQ holds $T_\Omega(x) = \mathcal{F}(x)$

NW 12.2

$c_1(x) = x_1^2 + x_2^2 - 2 = 0$

$c_1'(x) = (\qquad)^2 = 0$

$T_\Omega(x)$

$\begin{pmatrix} 0 \\ d_2 \end{pmatrix}$

LICQ

$\mathcal{F}(x) = \begin{pmatrix} 0 \\ d_2 \end{pmatrix}$

↳ $\nabla c_1 = \begin{pmatrix} -2\sqrt{2} \\ 0 \end{pmatrix}$

$\mathcal{F}(x) = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$ × LICQ.

$\nabla c_1' = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\sum_{i=1}^{N} \alpha_i v_i = 0$

$\alpha_1 v_1 = 0$

only $\alpha_1 = 0$

↳ First order Necessary Conditions (FONC)

Karush Kuhn Tucker (KKT) Conditions.

As before $\mathcal{L}(x,\lambda) = f(x) - \sum_i \lambda_i C_i(x)$

$i \in \mathcal{E} \cup \mathcal{I}$.

If

① If x^* is a local soln to $\min_{x \in \Omega} f(x)$.

② f and C_i's are continuously differentiable

③ LICQ holds at x^*

Then: There is a set of Lagrange multipliers s.t. the foll are satisfied at (x^*, λ^*):

a) $\nabla_x \mathcal{L}(x^*, \lambda^*) = 0$

b) $C_i(x^*) = 0$, $i \in \mathcal{E}$

c) $C_i(x^*) \geq 0$, $i \in \mathcal{I}$

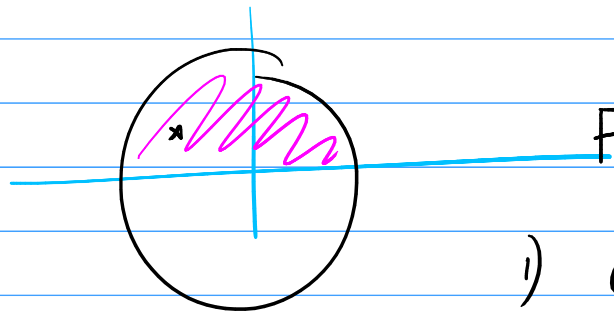
d) $\lambda_i \geq 0$, $i \in \mathcal{I}$

e) $\lambda_i C_i = 0$, $i \in \mathcal{E} \cup \mathcal{I}$

There can be many pts where a–e hold. But if ③ holds λ^* is unique.

↳ Proof sketch. $f(x) = x_1 + x_2$, s.t. $2 - x_1^2 - x_2^2 \geq 0$

$x_2 \geq 0$



FONC for x^* to be a local min

1) $c_1(x^*) \geq 0$, $c_2(x^*) \geq 0$

2) a) If $c_1(x^*) = 0$, $c_2(x^*) = 0$, then

$$\nabla_x \mathcal{L} = \nabla f(x^*) - \lambda_1 \nabla c_1 - \lambda_2 \nabla c_2 = 0, \quad \lambda_1 \geq 0, \lambda_2 \geq 0$$

b) $c_1(x^*) > 0$, $c_2(x^*) = 0$, then $\Rightarrow \lambda_1 = 0$

$$\nabla_x \mathcal{L} = \nabla f(x^*) - \lambda_2 \nabla c_2(x^*) = 0, \quad \lambda_2 \geq 0.$$

c) $c_1(x^*) = 0$, $c_2(x^*) > 0$, then $\Rightarrow \lambda_2 = 0$

$$\nabla_x \mathcal{L} = \nabla f(x^*) - \lambda_1 \nabla c_1(x^*) = 0, \quad \lambda_1 \geq 0$$

d) $c_1(x^*) > 0$, $c_2(x^*) > 0$, then $\lambda_1 = 0$, $\lambda_2 = 0$

$$\nabla_x \mathcal{L} = \nabla f(x^*) = 0.$$