Recap: OP, single inequality constraint:

$$\min_{\underline{x} \in \mathbb{R}^n} f(\underline{x}) \quad , \quad \text{subject to} \quad c_1(\underline{x}) \geq 0 .$$

Last week, we showed rigorously, that at the minimum $\underline{x}_*$, there exists a non-negative scalar $\lambda_1^* \geq 0$ such that:

$$\mathcal{L}(\underline{x}, \lambda_1) = f(\underline{x}) - \lambda_1 c_1(\underline{x})$$

$$\left\{ \begin{array}{c} \underline{\nabla} \mathcal{L}(\underline{x}_*, \lambda_1^*) = 0 \\ c_1(\underline{x}_*) \geq 0 \\ \lambda_1^* \geq 0 \\ \lambda_1^* c_1(\underline{x}_*) = 0 \end{array} \right\}$$

These are the KKT conditions for a single inequality constraint.

Aim of module: Derive the KKT conditions for a mixture of equality and inequality constraints. Along the way (this week) we introduce:

- The tangent cone
- Linearized feasible descent directions

This morning:

- State KKT conditions for pure inequality constraints
- Worked example

$\underline{CP}$: $\min_{x \in \mathbb{R}^n} f(\underline{x})$ subject to $c_i(\underline{x}) \geq 0$, $i \in \underbrace{\{1, 2, \ldots, m\}}_{\mathcal{I}}$

$\underline{\text{CP}}$: $\min_{\underline{x} \in \mathbb{R}^n} f(\underline{x})$ subject to $c_i(\underline{x}) \geq 0$, $i \in \underbrace{\{1, 2, \ldots, m\}}_{\mathcal{I}}$

No equality constraints: $\mathcal{E} = \emptyset$

If $\underline{x}_*$ is a minimizer, then there exist $\lambda_1^*, \ldots, \lambda_m^*$ such that:

$$\left\{ \begin{array}{l} \mathcal{L}(\underline{x}, \underline{\lambda}) = f(\underline{x}) - \lambda_1 c_1(\underline{x}) - \cdots - \lambda_m c_m(\underline{x}) \\ \underline{\lambda} = (\lambda_1, \ldots, \lambda_m)^T \end{array} \right\}$$

$$\left\{ \begin{array}{ll} \nabla \mathcal{L}(\underline{x}_*, \underline{\lambda}_*) & = 0 \\ c_i(\underline{x}_*) & \geq 0 \\ \lambda_i^* & \geq 0 \\ \lambda_i^* c_i(\underline{x}_*) & = 0 \end{array} \right\} \text{ KKT conditions}$$

Example (§ 13.3 — § 13.3.1, p. 108)

OP: $\min_{\underline{x} \in \mathbb{R}^2} f(\underline{x}) = x + y$

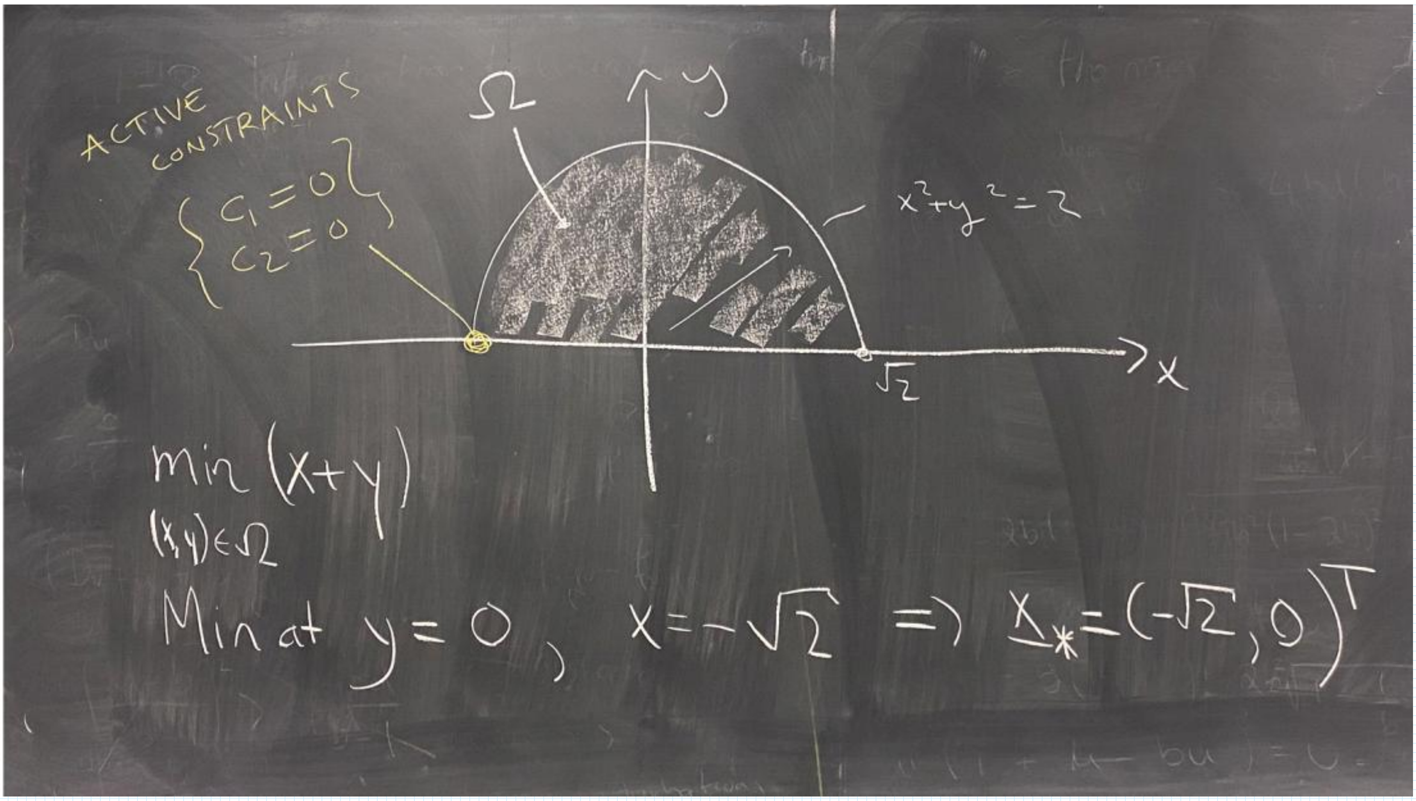
subject to $c_1(\underline{x}) \geq 0$, $c_2(\underline{x}) \geq 0$

where $c_1(\underline{x}) = 2 - x^2 - y^2$

$c_2(\underline{x}) = y$

Feasible set: $y \geq 0 \Rightarrow$ upper half-plane

$2 - x^2 - y^2 \geq 0 \Rightarrow x^2 + y^2 \leq 2 \Rightarrow$ inside disc of radius $\sqrt{2}$.



$$\mathcal{L}(\underline{x},\underline{\lambda}) = f(\underline{x}) - \lambda_1 c_1(\underline{x}) - \lambda_2 c_2(\underline{x})$$

$$= (x+y) - \lambda_1(2-x^2-y^2) - \lambda_2 y$$

$$\nabla\mathcal{L} = \underline{i} + \lambda_1 \cdot 2x\,\underline{i}$$
$$+\underline{j} + \lambda_1 \cdot 2y\,\underline{j} - \lambda_2 \underline{j}$$

$\nabla\mathcal{L}=0$ @ minimizer :

$$1+2\lambda_1^* x_* = 0$$
$$1+2\lambda_1^* y_* = \lambda_2^*$$

Omit stars from now on:

$$\left\{\begin{array}{l} 1+2\lambda_1 x = 0 \\ \boxed{1+2\lambda_1 y = \lambda_2} \end{array}\right.$$

– "First gradient condition," $\frac{\partial\mathcal{L}}{\partial x}=0$

– "Second gradient condition", $\frac{\partial\mathcal{L}}{\partial y}=0$

Solve 4 equations

$$\boxed{1 + 2\lambda_1 y = \lambda_2}$$ — "Second gradient condition", $\frac{\partial \mathcal{L}}{\partial y} = 0$

Complementarity:

$$\lambda_1 (2 - x^2 - y^2) = 0, \quad \lambda_2 \cdot y = 0$$

Solve 4 equations

Take eq^{n} in box and multiply by λ_2:

$$\lambda_2 + 2\lambda_1 \cancel{(\lambda_2 y)} = \lambda_2^2$$

complementarity condition

$$\Rightarrow \lambda_2 = \lambda_2^2 \Rightarrow \begin{cases} \lambda_2 = 1 \\ \lambda_2 = 0 \end{cases}$$

Case 1. Look at $\lambda_2 = 0$. Back to eq^{n} in box:

$$1 + 2\lambda_1 y = \cancel{\lambda_2}$$

$$\Rightarrow \lambda_1 y = -1/2$$

$\Rightarrow$ λ_1 or y is negative.

But $\lambda_1 \geq 0$ (by KKT) and $y \geq 0$ (feasibility)

Hence, Case 1 is impossible.

Case 2: We are forced into case 2, with $\lambda_2 = 1$.

Back to eq^{n} in box:

$$1 + 2\lambda_1 y = 1$$

$$\Rightarrow \lambda_1 y = 0$$

Back to first gradient condition:

Back to first gradient condition:

$$1 + 2\lambda_1 x = 0$$

Hence, λ_1 can't be zero (because $1 = 0$ is a false statement)

Hence, $\lambda_1 > 0$ (C_1 is an active constraint)

But $\lambda_1 y = 0$, hence $y = 0$ (C_2 is an active constraint)

Both constraints are active — in agreement with figure.

Wrap up:

$$\underbrace{\lambda_1}_{\neq 0} (2 - x^2 - \cancel{y^2}) = 0 \qquad \text{COMPLEMENTARITY}$$

$$\Rightarrow \quad x = \pm\sqrt{2}.$$

$x_* = -\sqrt{2}$ is the minimizer.

$y_* = 0$

Lagrange multipliers:

First gradient condition: $1 + 2\lambda_1 x = 0 \Rightarrow \lambda_1^* = + \frac{1}{2\sqrt{2}}$

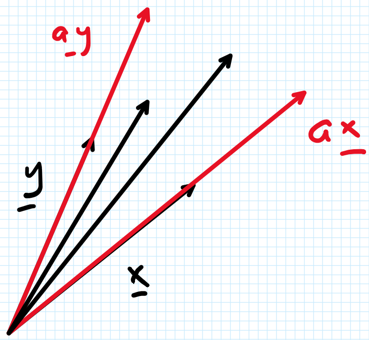
Previously worked out: $\lambda_2^* = 1$

Chapter 14 — Tangent Cone

§ 14.1 A cone in $\mathbb{R}^n$

§ 14.1 A cone in $\mathbb{R}^n$

Def: A set $\mathcal{C}$ in $\mathbb{R}^n$ is a cone, if for each $\underline{x} \in \mathcal{C}$, the vector $a\underline{x}$ is also in $\mathcal{C}$, where a is a positive constant.



__Example__: Let $\underline{a}_1, \dots, \underline{a}_m$ be vectors in $\mathbb{R}^n$. Then the set

$$\mathcal{C} = \{ \underline{a}_1 x_1 + \dots + \underline{a}_m x_m \mid x_i \geq 0, \ i = 1, 2, \dots, m \}$$

is a cone.

Introduce the matrix

$$A = \begin{array}{c} \begin{pmatrix} | & & | \\ \underline{a}_1 & \cdots & \underline{a}_m \\ | & & | \end{pmatrix} \\ \longleftarrow m \text{ columns} \longrightarrow \end{array} \begin{array}{c} \uparrow \\ n \text{ rows} \\ \downarrow \end{array} \in \mathbb{R}^{n \times m}$$

$$\mathcal{C} = \{ \underline{a}_1 x_1 + \dots + \underline{a}_m x_m \mid x_i \geq 0, \ i = 1, \dots, m \}$$

$$= \{ A\underline{x} \mid \underline{x} \in \mathbb{R}^m, \ x_i \geq 0 \}$$

$$= \mathcal{C}(A)$$

Double-check:

$$(A\underline{x})_i = \sum_j A_{ij} x_j$$

$$= A_{i1} x_1 + \cdots + A_{im} x_m$$

$$A\underline{x} = \begin{pmatrix} A_{11} \\ A_{21} \\ \vdots \\ A_{n1} \end{pmatrix} x_1 + \cdots + \begin{pmatrix} A_{1m} \\ A_{2m} \\ \vdots \\ A_{nm} \end{pmatrix} x_m$$

$$= \underline{a}_1 x_1 + \cdots + \underline{a}_m x_m$$

Tangent cone (§ 14.2)

OP

$$\min_{\underline{x} \in \mathbb{R}^n} f(\underline{x}) \quad \text{subject to} \quad \begin{cases} c_i(\underline{x}) = 0 & i \in \mathcal{E} \\ c_i(\underline{x}) \geq 0 & i \in \mathcal{I} \end{cases}$$

Feasible set:

$$\Omega = \left\{ \underline{x} \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} c_i(\underline{x}) = 0, & i \in \mathcal{E} \\ c_i(\underline{x}) \geq 0, & i \in \mathcal{I} \end{array} \right\}$$

Def^{n}: The vector $\underline{d}$ is a tangent vector to Ω at the point $\underline{x} \in \Omega$ if there is a feasible sequence $\{\underline{z}_k\}_{k=0}^{\infty}$ approaching $\underline{x}$ and a sequence of positive scalars $\{t_k\}_{k=0}^{\infty}$ such that $t_k \to 0$ and $k \to \infty$,

$$\underline{d} = \lim_{k\to\infty} \frac{\underline{z}_k - \underline{x}}{t_k}.$$

The set of all such tangent vectors $\underline{d}$ is called the tangent cone at $\underline{x}$, $T_\Omega(\underline{x})$.

We double-check that $T_\Omega(\underline{x})$ is a cone.

Take $\underline{d} \in T_\Omega(\underline{x})$. There exists $\{\underline{z}_k\}_{k=0}^{\infty}$ feasible and $\{t_k\}_{k=0}^{\infty}$ with $t_k > 0$ and $t_k \to 0$ as $k \to \infty$ such that:

$$\underline{d} = \lim_{k\to\infty} \frac{\underline{z}_k - \underline{x}}{t_k}$$

Take $\alpha > 0$, and

$$\alpha \underline{d} = \alpha\left(\lim_{k\to\infty} \frac{\underline{z}_k - \underline{x}}{t_k}\right)$$

$$= \lim_{k \to \infty} \frac{\underline{z}_k - \underline{x}}{\frac{1}{\alpha} t_k}$$

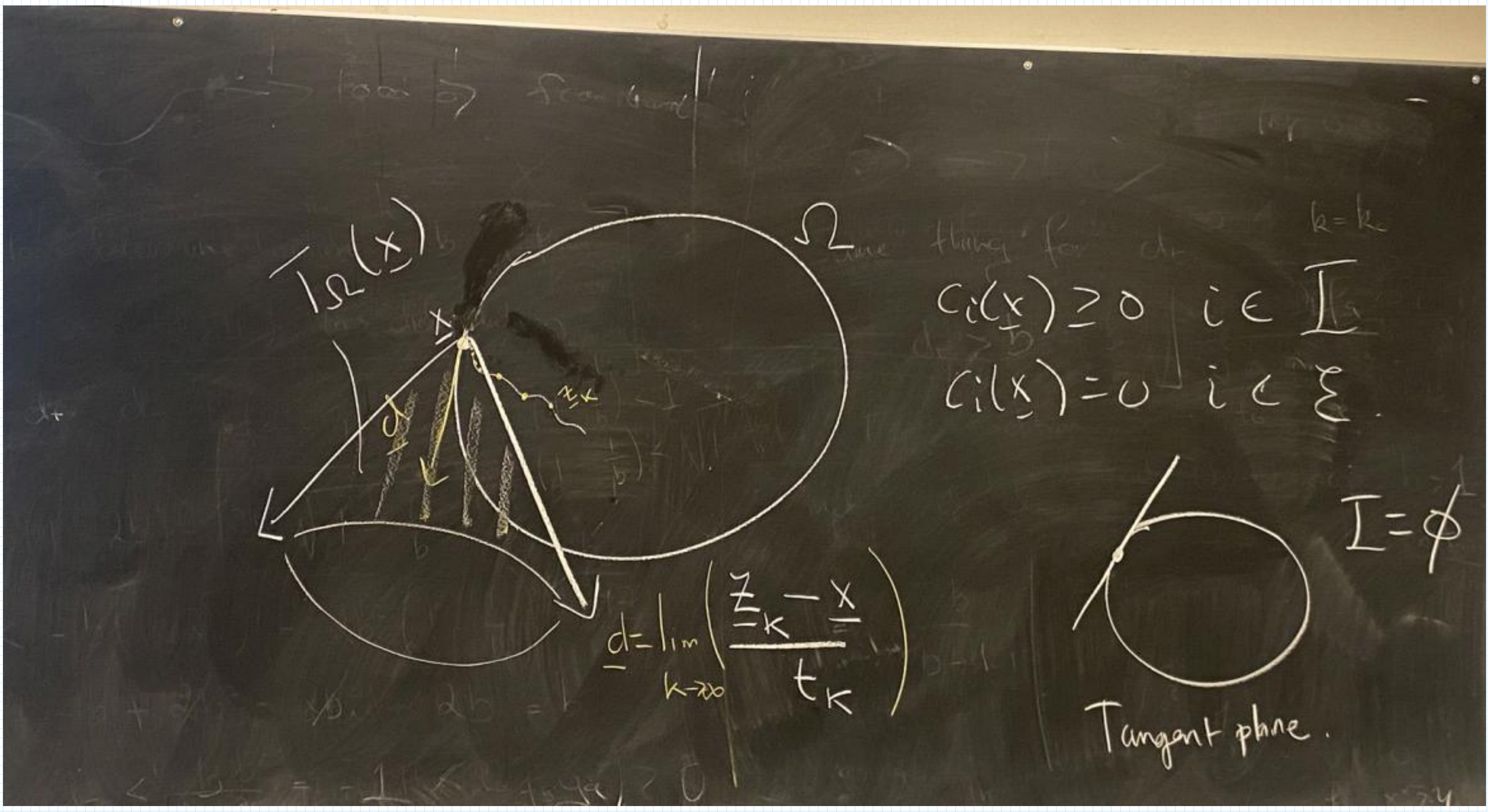
Feasible sequences for $\alpha \underline{d}$:

- $\{\underline{z}_k\}_{k=0}^{\infty}$
- $\{\frac{1}{\alpha} t_k\}_{k=0}^{\infty} = \{\tilde{t}_k\}_{k=0}^{\infty}$

$$\alpha \underline{d} = \lim_{k \to \infty} \frac{\underline{z}_k - \underline{x}}{\tilde{t}_k} \in T_\Omega(\underline{x}).$$

Hence: $\underline{d} \in T_\Omega(\underline{x}) \Rightarrow \alpha \underline{d} \in T_\Omega(\underline{x}), \alpha > 0$

$\Rightarrow T_\Omega(\underline{x})$ is a cone in $\mathbb{R}^n$.



OP:

$$\min_{\underline{x} \in \mathbb{R}^n} f(\underline{x}), \quad \text{subject to} \quad \begin{cases} c_i(\underline{x}) = 0, & i \in \mathcal{E} \\ c_i(\underline{x}) \geq 0, & i \in \mathcal{I} \end{cases}$$

Terminology. $\begin{cases} i \in \mathcal{I}, & c_i(\underline{x}) = 0 \text{ is an active constraint} \\ i \in \mathcal{I}, & c_i(\underline{x}) > 0 \text{ is an inactive constraint} \end{cases}$

KKT conditions:

$$\mathcal{L}(\underline{x}, \underline{\lambda}) = f(\underline{x}) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} c_i(\underline{x}) \lambda_i, \quad \underline{\lambda} = (\lambda_1, \dots, \lambda_m)^T$$

At the minimizer $\underline{x}_*$, there exists $\underline{\lambda}_*$ such that:

$$\left.\begin{aligned} &\nabla \mathcal{L}(\underline{x}_*, \underline{\lambda}_*) = 0 \\ &c_i(\underline{x}_*) \geq 0 \qquad i \in \mathcal{I} \\ &c_i(\underline{x}_*) = 0 \qquad i \in \mathcal{E} \\ &\lambda_i^* \geq 0 \qquad i \in \mathcal{I} \\ &\lambda_i^* c_i(\underline{x}_*) = 0 \end{aligned}\right\} \text{KKT conditions}$$

Already:

- Proved KKT conditions rigorously for a single equality constraint
- " " " " for a single inequality constraint.

Building up now to general case.

Building up now to general case.

To do this, we need to define the tangent cone:

- $\underline{x} \in \Omega$ feasible. Let $\{\underline{z}_k\}_{k=0}^{\infty}$ be a feasible sequence with $\underline{z}_k \to \underline{x}$ as $k \to \infty$. Let $\{t_k\}_{k=0}^{\infty}$ with $t_k > 0$ and $t_k \to 0$ as $k \to \infty$.

  Then

$$\underline{d} = \lim_{k \to \infty} \frac{\underline{z}_k - \underline{x}}{t_k} \in T_\Omega(\underline{x}).$$

Today: The Linearized Feasible Descent Directions (§14.2, p. 113)

(LFDDs):

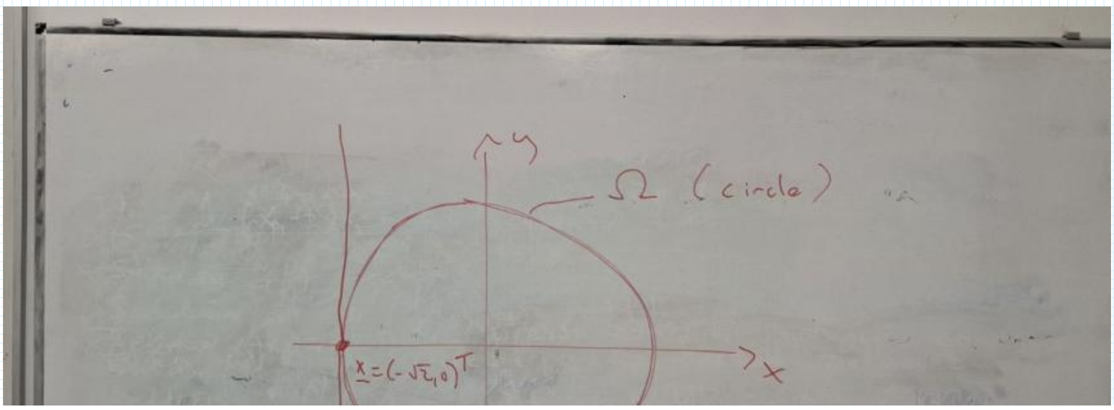
$$\mathcal{F}(\underline{x}) = \left\{ \underline{d} \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} \underline{d} \cdot \underline{\nabla} c_i(\underline{x}) = 0 & i \in \mathcal{E} \\ \underline{d} \cdot \underline{\nabla} c_i(\underline{x}) \geq 0 & i \in \mathcal{I} \end{array} \right\}$$

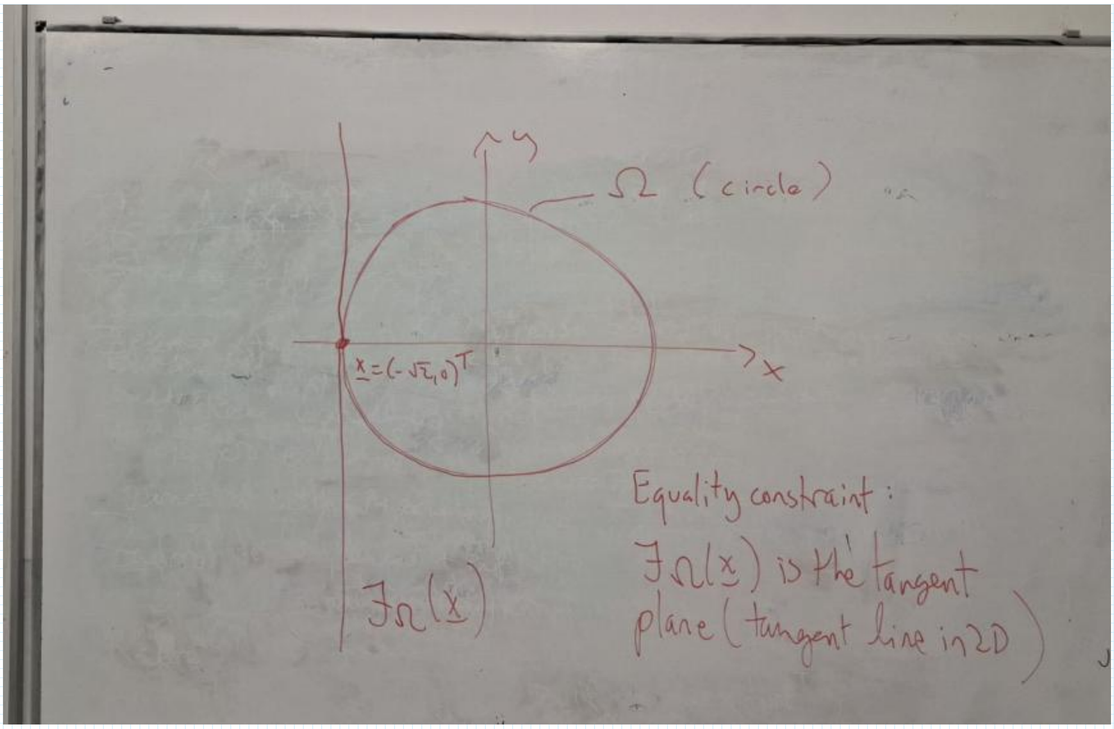
Example:

$$\min f(\underline{x}) = x + y$$

subject to $c_1(\underline{x}) = 0$, $c_1(\underline{x}) = 2 - x^2 - y^2$.

We compute the LFDDs at $\underline{x} = (-\sqrt{2}, 0)^T$.





$$\nabla c_1 = -2x\,\underline{i} - 2y\,\underline{j}$$

$$\nabla c_1(\underline{x}) = -2(-\sqrt{2})\,\underline{i} - 2(0)\,\underline{j} = 2\sqrt{2}\,\underline{i}$$

$$\mathcal{F}(\underline{x}) = \left\{ \underline{d} = (d_1, d_2)^T \;\middle|\; \underline{d} \cdot \nabla c_1(\underline{x}) = 0 \right\}$$

$$(d_1, d_2)^T \cdot 2\sqrt{2}\,\underline{i} = 0 \implies d_1 = 0$$

No condition on d_2.

$$\therefore\ \mathcal{F}_\Omega(\underline{x}) = \left\{ (0, d_2) \;\middle|\; d_2 \in \mathbb{R} \right\}$$

i.e. the y-axis.

Tangent cone: Construct a feasible sequence approaching $\underline{x}$.

$$\underline{x}_k = \left(\sqrt{2}\cos(\pi - \theta_k) \,,\ \pm\sqrt{2}\sin(\pi - \theta_k) \right)^T$$

$$\underline{x}_k = \left(\sqrt{2}\cos(\pi-\theta_k) \,,\ \pm\sqrt{2}\sin(\pi-\theta_k) \right)'$$

Take $\theta_k \to 0$ as $k \to \infty$.

Then $\underline{x}_k \longrightarrow \left(\sqrt{2}(-1) \,,\ 0 \right)^T$ ☑

$$\underline{x}_k - \underline{x} = \left(\sqrt{2}\cos(\pi-\theta_k) - (-\sqrt{2}) \,,\ \pm\sqrt{2}\sin(\pi-\theta_k) \right)^T$$

$$= \begin{pmatrix} \sqrt{2}\,\cancel{\cos(\pi)} + 0 + (+1)\frac{1}{2}\theta_k^2 + O(\theta_k^3) \cancel{+\sqrt{2}} \,, \\ \pm\sqrt{2}\,\cancel{\sin(\pi)} \pm \sqrt{2}\cos(\pi)(-\theta_k) + O(\theta_k^2) \end{pmatrix}^T$$

$$= \left(\tfrac{1}{2}\theta_k^2 + O(\theta_k^3) \,,\ \pm\sqrt{2}\,\theta_k + O(\theta_k^2) \right)^T$$

$$\frac{\underline{x}_k - \underline{x}}{\theta_k} = \left(\tfrac{1}{2}\theta_k + O(\theta_k^2) \,,\ \pm\sqrt{2}\,\cancel{\theta_k} + O(\theta_k) \right)^T$$

$$\lim_{\theta_k \to 0} \frac{\underline{x}_k - \underline{x}}{\theta_k} = (0, +\sqrt{2})^T \text{ or } (0, -\sqrt{2})^T$$

Hence, $\underline{d} = \left\{ \begin{matrix} (0, \sqrt{2})^T \\ (0, -\sqrt{2})^T \end{matrix} \right\} \in T_\Omega(\underline{x})$

$T_\Omega(\underline{x})$ closed under $\underline{x} \to \alpha\underline{x}$, $\alpha > 0$.

Hence,

$$T_\Omega(\underline{x}) = \left\{ (0, d_2) \mid d_2 \in \mathbb{R} \right\}$$

$$= \mathcal{F}_\Omega(\underline{x})$$

In this case, $\boxed{T_\Omega(\underline{x}) = \mathcal{F}_\Omega(\underline{x})}$

Nice property to have, but not guaranteed.

E.g. $c_1(\underline{x}) = (2 - x^2 - y^2)^2$

$$c_1(\underline{x}) = 0.$$

Same picture as before!

$$\nabla c_1 : \begin{cases} 2(\cancel{2 - x^2 - y^2})(-2x) & x\text{-dir}^n \\ 2(\cancel{2 - x^2 - y^2})(-2y) & y\text{-dir}^n \end{cases}$$

Evaluate at any feasible point:

$$\boxed{\nabla c_1 = 0}.$$

$$\mathcal{F}_\Omega(\underline{x}) = \{ \underline{d} \in \mathbb{R}^2 \mid \underbrace{\underline{d} \cdot \nabla c_1}_{=0} = 0 \}$$

No restriction on $\underline{d}$ any more.

$$\mathcal{F}_\Omega(\underline{x}) = \mathbb{R}^2.$$

$$T_\Omega(\underline{x}) = \{ (0, d_2) \mid d_2 \in \mathbb{R} \}$$

Here, $T_\Omega(\underline{x}) \neq \mathcal{F}_\Omega(\underline{x})$.

Generalizing, we can guess that

$$T_\Omega(\underline{x}) \neq \mathcal{F}_\Omega(\underline{x})$$

when the ∇c_i's are not all linearly independent (we will show this).

<u>Next</u>: We will look at the same example but with an inequality constraint (p. 115).

$$\min_{\underline{x} \in \mathbb{R}^2} f(\underline{x}) = x + y \quad \text{subject to} \quad c_1(\underline{x}) \geq 0$$

where $c_1(\underline{x}) = 2 - x^2 - y^2$.

LFDDs (easier), computed at $\underline{x} = (-\sqrt{2}, 0)^T$.

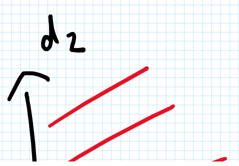
$$\nabla c_1(\underline{x}) = 2\sqrt{2}\,\underline{i} + 0\,\underline{j}$$

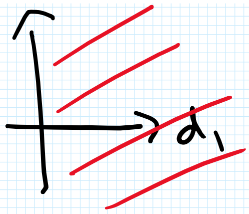
$$\underline{d} \cdot \nabla c_1(\underline{x}) \geq 0$$

$$\Rightarrow \quad d_1 (2\sqrt{2}) \geq 0 \quad \Rightarrow \quad d_1 \geq 0.$$

No restriction on d_2.

$$\mathcal{F}_\Omega(\underline{x}) = \{ (d_1, d_2)^T \mid d_1 \geq 0 \}.$$





Tangent cone:

$$\underline{d} = \lim_{k \to \infty} \frac{\underline{x}_k - \underline{x}}{t_k}$$

Re-arrange:

$$\underline{x}_k = \underline{x} + t_k \underline{d} + \underline{\epsilon}_k$$

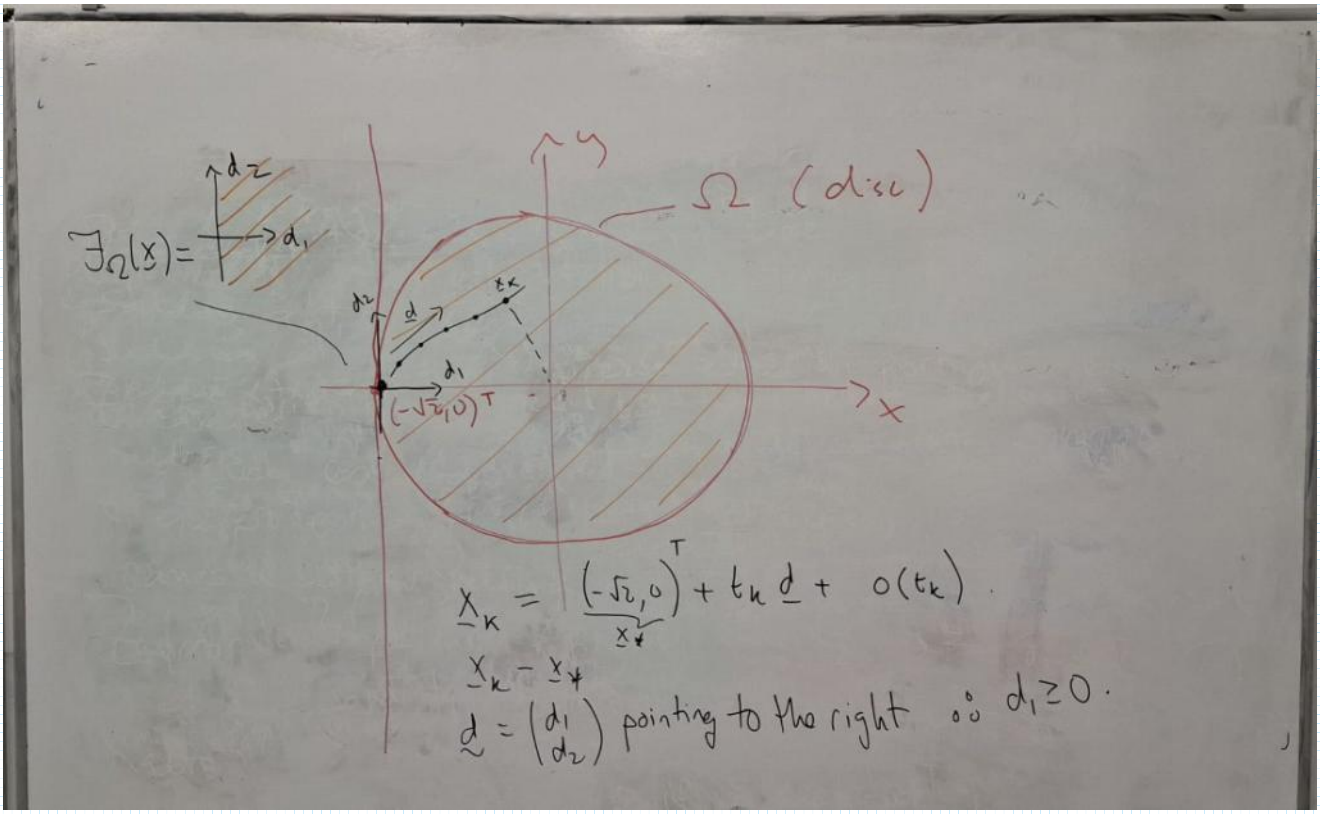
Re-arrange again:

$$\frac{\underline{x}_k - \underline{x}}{t_k} - \frac{\underline{\epsilon}_k}{t_k} = \underline{d}$$

So we require $\boxed{\epsilon_k / t_k \longrightarrow 0 \quad \text{as} \quad t_k \to 0.}$

We say: "$\underline{\epsilon}_k$ is little-o of t_k": $\underline{\epsilon}_k = o(t_k)$.

$$\underline{x}_k = (-\sqrt{2}, 0)^T + t_k \underline{d} + o(t_k)$$



Referring to the figure, $\underline{d}$ is pointing to the right, giving $d_1 \geq 0$ and no restriction on d_2, hence:

$$T_\Omega(\underline{x}) = \{(d_1, d_2) \mid d_1 \geq 0\}$$

$$= \mathcal{F}_\Omega(\underline{x}) \qquad \boxed{\checkmark}.$$

In general, a sufficient condition that guarantees

$$T_\Omega(\underline{x}) = \mathcal{F}_\Omega(\underline{x})$$

is the LICQ.

Def^n (14.3)

Def$^{\underline{n}}$ (14.3)

LICQ / Linear Independence constraint qualification.

Given the feasible point $\underline{x}$ and the active set $\mathcal{A}(\underline{x})$ we say that the LICQ holds at $\underline{x}$ if the active constraints

$$\nabla c_i(\underline{x}), \quad i \in \mathcal{A}(\underline{x})$$

are linearly independent.

We will show: If the LICQ hold at $\underline{x}$, then

$$T_\Omega(\underline{x}) = \mathcal{F}_\Omega(\underline{x}).$$

Examples we looked at today:

- One constraint.

LICQ holds (trivially)

- $\nabla c_1(\underline{x}) \neq 0, \quad \mathcal{F}_\Omega(\underline{x}) = T_\Omega(\underline{x})$

LICQ does not hold

- $\nabla c_1(\underline{x}) = 0, \quad \mathcal{F}_\Omega(\underline{x}) \neq T_\Omega(\underline{x})$

Optimization Algorithms
(ACM 41030)

Dr Lennon Ó Náraigh

Exercises #5

1. Does the OP

$$\min f(\boldsymbol{x}) = (y+100)^2 + \tfrac{1}{100}x^2$$

subject to $y - \cos x \geq 0$ have a finite or infinite number of local solutions? Use the KKT conditions to justify your answer.

KKT conditions for a single inequality constraint.

KKT conditions for a single inequality constraint.

At the minimizer $\underline{x}_*$, there exists $\lambda_1^* \geq 0$ such that:

$$\nabla \mathcal{L}(\underline{x}_*, \lambda_1^*) = 0$$
$$c_1(\underline{x}_*) \geq 0$$
$$\lambda_1^* \geq 0$$
$$\lambda_1^* c_1(\underline{x}_*) = 0$$

Apply to:

$$\lambda_1 \underline{\cos(x)}$$

$$\mathcal{L} = (y+100)^2 + \frac{1}{100}x^2 - \lambda_1\left(y - \cos(x)\right)$$

$$c_1: \quad y - \cos(x) \geq 0.$$

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{1}{100}\cdot 2x - \underline{\lambda_1 \sin(x)}$$

$$\frac{\partial \mathcal{L}}{\partial y} = 2(y+100) - \lambda_1 \underline{1}$$

Omit "*" notation for minimizer.

$$\frac{\partial \mathcal{L}}{\partial x} = 0 \Rightarrow \frac{x}{50} - \lambda_1 \sin(x) = 0 \Rightarrow \boxed{\frac{x}{50} = \lambda_1 \sin(x)} \quad (1)$$

$$\frac{\partial \mathcal{L}}{\partial y} = 0 \Rightarrow \boxed{2(y+100) = \lambda_1} \quad (2)$$

Solve for x, y, and λ_1 (three equations).

Complementarity condition gives third equation:

complementarity condition gives third equation:

$$\boxed{\lambda_1 (y - \cos(x)) = 0} \quad (3)$$

Case 1: $\lambda_1 = 0$, C.C. (3) OK.

Back to (2): $y = -100$ NOT FEASIBLE, since $y \geq \cos(x)$.

So we are forced into Case 2.

Case 2: $\lambda_1 \neq 0$, $\lambda_1 > 0$.

By the C.C. (3) we have $y = \cos(x)$

So the constraint is active ($c_1(\underline{x}) = 0$).

Back to (1) and (2):

$$\frac{1}{50} x = \lambda \sin(x)$$

$$2(y+100) = \lambda \quad , \quad \lambda \neq 0.$$

Since the second eqⁿ here is never zero, we can divide one by the other to get:

$$\frac{\frac{1}{50} x}{2(y+100)} = \frac{\cancel{\lambda} \sin(x)}{\cancel{\lambda}}$$

$$\Rightarrow \frac{1}{50} x = 2 \sin(x) [y + 100]$$

$$\Rightarrow \frac{1}{100} x = \sin(x) (y + 100)$$

Active constraint, $y = \cos(x)$.

Active constraint, $y = \cos(x)$.

$$\boxed{\frac{1}{100} x = \sin(x)\left(100 + \cos(x)\right)} \qquad (4)$$

The minimizer x_* solves the rootfinding condition (4).

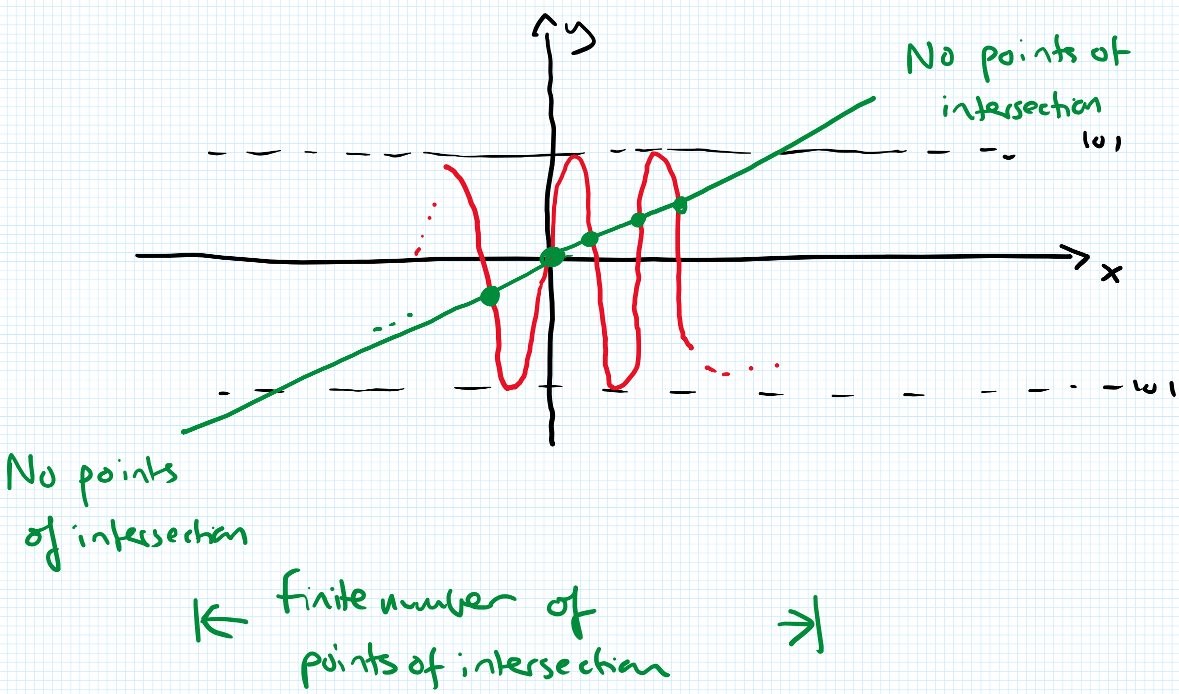
Finitely many roots or infinitely many roots?

Look at (4) again:

$$\underbrace{\frac{1}{100} x}_{y_1(x)} = \underbrace{\sin(x)\left(100 + \cos(x)\right)}_{y_2(x)}$$

Roots of (4) are the points of intersection of two curves.

- Curve y_1 is unbounded: $|y_1| \to \infty$ as $|x| \to \infty$.
- Curve $|y_2| \leq 1 \times 101$ bounded.



This corrects Week 8, Lecture 1, where I said this OP has an infinite number of local minima. Instead, the OP has a large (but finite) number of local minima.

We can also find the answer without resorting to the KKT conditions, so long as we can accept that the constraint is active.

Active constraint, $y = \cos(x)$. We can reparametrize the cost function:

$$f(x, y = \cos(x)) = \left(\cos(x) + 100\right)^2 + \frac{1}{100} x^2 = \tilde{f}(x)$$

We compute the minimizer of $\tilde{f}(x)$ by looking at

$$\frac{d\tilde{f}}{dx} = 0$$

For large x, $\tilde{f}(x) \sim \frac{1}{100} x^2$

So for large x, $\frac{d\tilde{f}(x)}{dx} \sim \frac{1}{50} x \neq 0$ (large $|x|$).

Hence,

Zeros are contained in a finite interval $(-R, R)$.

Hence,

Since $\tilde{f}$ is a smooth function there can only

Hence,

Since $\bar{F}$ is a smooth function, there can only be finitely many zeros.

4. Consider the OP

$c_1(\underline{x})$ CIRCLE

$$\min(x+y), \qquad \text{subject to } 2 - x^2 - y^2 = 0.$$

Specify two feasible sequences that approach the **maximizing** point $(1,1)^T$ and show that neither sequence is a decreasing sequence for f.

$\underline{x}_* = (1,1)$ maximizes the cost function.

Because the constraint is simple, we may reparametrize the feasible set using polar coordinates:

$$\underline{\Omega} = \left\{ \underline{x} = \sqrt{2}\cos\left(\frac{\pi}{4}+\phi\right)\underline{i} + \sqrt{2}\sin\left(\frac{\pi}{4}+\phi\right)\underline{j} \,\middle|\, \phi \in [0,2\pi) - \frac{\pi}{4} \right\}$$

Let

$$\underline{x}(\phi) = \sqrt{2}\cos\left(\frac{\pi}{4}+\phi\right)\underline{i} + \sqrt{2}\sin\left(\frac{\pi}{4}+\phi\right)\underline{j}$$

Hence,

$$\underline{x}(0) = \sqrt{2}\frac{1}{\sqrt{2}}\underline{i} + \sqrt{2}\frac{1}{\sqrt{2}}\underline{j} = \underline{i}+\underline{j} = \underline{x}_*$$

Re-parametrize the cost function:

$$f(\underline{x}) = f(\underline{x}(\phi))$$

$$= x+y$$

$$= \sqrt{2}\cos\left(\frac{\pi}{4}+\phi\right) + \sqrt{2}\sin\left(\frac{\pi}{4}+\phi\right)$$

$$= \sqrt{2}\left(\cos\left(\frac{\pi}{4}+\phi\right) + \sin\left(\frac{\pi}{4}+\phi\right)\right)$$

$$= \tilde{f}(\phi)$$

because we are given

$$= \tilde{f}(\phi)$$

because we are given that $\underline{x}_*$ is a maximizer.

We are interested in the point $\underline{x}_*$ and in constructing a feasible non-decreasing sequence tending to $\underline{x}_*$.

$$\tilde{f}(\phi) = \tilde{f}(0) + \phi \left.\frac{d\tilde{f}}{d\phi}\right|_0 + \frac{1}{2}\left.\frac{d^2\tilde{f}}{d\phi^2}\right|_0 \phi^2 + \text{H.O.T.}$$

$$\tilde{f}(\phi) = \sqrt{2}\left(\cos\left(\frac{\pi}{4}+\phi\right) + \sin\left(\frac{\pi}{4}+\phi\right)\right)$$

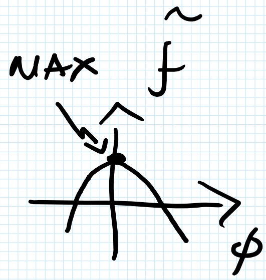
$$\tilde{f}(0) = 2 .$$

$$\left.\frac{d\tilde{f}}{d\phi}\right|_0 = 0$$

$$\left.\frac{d^2\tilde{f}}{d\phi^2}\right|_0 = -2 \qquad 2 \times \frac{1}{2}$$

By the Taylor expansion,

$$\tilde{f}(\phi) = 2 - \phi^2 + \text{H.O.T.}$$



We need to construct a feasible sequence tending to $\underline{x}_*$.

Solution

$$\underline{x}_n = \underline{x}(\phi_n), \quad \phi_n = \begin{cases} \frac{1}{n}, & n \geq 1 \end{cases}$$

$$\underline{x}_n = \underline{x}(\phi_n) \quad , \quad \phi_n = \begin{cases} n \\ -\frac{1}{n} \quad , n \geq 1 \end{cases}$$

We check that the cost function is non-decreasing on those sequences (as $n \to \infty$).

$$f(\underline{x}_n) = \tilde{f}(\phi_n)$$

$$\overset{n\to\infty}{=} 2 - \phi_n^2$$

$$= 2 - \frac{1}{n^2}$$

$$f(\underline{x}_{n+1}) - f(\underline{x}_n) = \left(\cancel{2} - \frac{1}{(n+1)^2}\right) - \left(\cancel{2} - \frac{1}{n^2}\right)$$

$$= \frac{1}{n^2} - \frac{1}{(n+1)^2}$$

$$= \frac{(n+1)^2 - n^2}{n^2(n+1)^2}$$

$$= \frac{\cancel{n^2} + 2n + 1 - \cancel{n^2}}{n^2(n+1)^2} > 0$$

$$\therefore \; f(\underline{x}_{n+1}) - f(\underline{x}_n) > 0$$

I.e. the cost $f^{\underline{n}}$ is an increasing function along this sequence.

Remark: This is consistent with the point (1,1)

Remark: This is consistent with the point $(1,1)$ being a maximizer. $\blacksquare$