Theory of Constrained Optimization. Week 3

Note Title 1/25/2022

The KKT condition is insufficient for guaranteeing a local minimizer.

Well, we know that even for a 1-D unconstrained problem the necessary $f'(x^*)=0$ condition does not guarantee x^* is a local (let alone global) minimizer.

e.g. $f(x)=x^3$, $f'(0)=0$ but 0 is not a local minimizer.

Note: f is not convex and $f''(0)=0$ (2nd derivative test failed)

Two important results:

(1) If the optimization problem is convex, then the KKT condition is sufficient to guarantee that x^* is a global minimizer.

(2) For general (nonconvex) problems, an additional second order condition guarantees that x^* is a local minimizer. [Sec 12.5, N&W]

Convexity

Def: A set $C \subseteq \mathbb{R}^n$ is convex if : $\forall\, x, y \in C$, $t \in [0,1]$, $(1-t)x + ty \in C$.

Def: Let $C \subseteq \mathbb{R}^n$ be convex. A function $f : C \to \mathbb{R}$ is called convex if
$$f((1-t)x + ty) \leq (1-t) f(x) + t f(y), \ \forall\, x, y \in C,\ t \in [0,1].$$

f is called concave if $-f$ is convex.

Easy fact : intersection of convex sets is convex.

If $c_1, \cdots, c_m : \mathbb{R}^n \to \mathbb{R}$ are concave, then each $\{x \in \mathbb{R}^n : c_i(x) \geq 0\}$ is convex, and so is
$$\Omega = \{x : c_1(x) \geq 0, \cdots, c_m(x) \geq 0\} = \bigcap_{i=1}^{m} \{x \in \mathbb{R}^n : c_i(x) \geq 0\}.$$

[Note : $c : \mathbb{R}^n \to \mathbb{R}$ is convex $\not\Longrightarrow$ $\{x : c_i(x) \geq 0\}$ or $\{x : c_i(x) = 0\}$ is convex.]

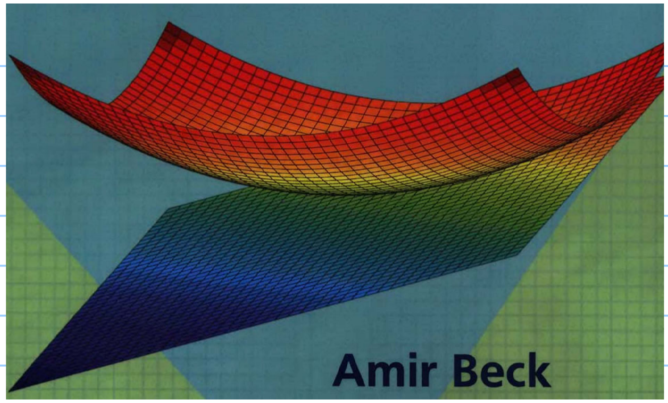
Another easy fact : Any affine function $f(x) = a^T x + b$, $a \in \mathbb{R}^n$, $b \in \mathbb{R}$ is both convex and concave.

Thm (the gradient inequality) a local property

Let $f:\mathbb{R}^n\to\mathbb{R}$ be C^1. $f|_C$ is convex. a global property

f is a convex $\iff f(x)+\nabla f(x)^T(y-x)\le f(y),\ \forall x,y\in C.$

The proof is not hard, see [Becks] Thm 7.6.



Thm. Let $f:\mathbb{R}^n\to\mathbb{R}$ be C^1. $f|_C$ is convex.

$$\nabla f(x^*)=0 \implies x^* \text{ is a global minimizer of } f: C\to\mathbb{R}.$$

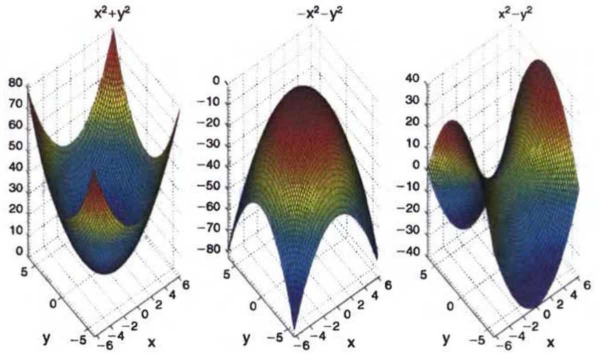
Converse is not true, and that is the whole point of KKT!

Proof: By the gradient inequality, $f(z)\ge f(x^*)+\underbrace{\nabla f(x^*)}_{=0}{}^T(z-x^*)$
$= f(x^*).$

QED.

Thm Same setting as above, but now f is C^2.

f is convex $\iff \nabla^2 f(x)\succeq 0 \quad \forall x\in C.$



Corollary: $f(x)=x^TAx+2b^Tx+c$ is convex $\iff A\succeq 0.$

a global property   a local property

KKT meets convexity

Let $f, g_1, \cdots, g_m : \mathbb{R}^n \to \mathbb{R}$ be convex C^1 functions.

$\Leftrightarrow -g_i(x) \geq 0$

Consider the convex optimization problem : $\min f(x)$ s.t. $g_i(x) \leq 0, \ i=1,\cdots,m$ $(*)$

<u>Thm</u> : The KKT conditions are satisfied at x^* $\Rightarrow$ x^* is a global minimizer of $(*)$.

Note : $(\Leftarrow)$ requires a constraint qualification condition
$(\Rightarrow)$ does not require CQ.

Proof : The KKT conditions for $(*)$ at x^* are $g_i(x^*) \leq 0$,

$$\underbrace{\nabla f(x^*) = \sum_{i=1}^m \lambda_i^* \nabla(-g_i)(x^*)}_{\Leftrightarrow\ \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) \ (= \nabla_x \mathcal{L}(x^*, \lambda^*)) \ = 0.} \quad , \quad \lambda_i^* \geq 0 \quad , \quad \lambda_i^* g_i(x^*) = 0.$$

The function $s(x) = \underset{\text{convex}}{f(x)} + \sum_i \underset{\geq 0}{\lambda_i^*} \underset{\text{convex}}{g_i(x)} \ (= \mathcal{L}(x, \lambda^*))$ is convex.

Also, $\nabla s(x^*) = \nabla f(x^*) + \sum_1^m \lambda_i^* \nabla g_i(x^*) \overset{\text{KKT}}{=} 0$

So, by convexity, x^* is a global minimizer of $s(\cdot)$ over $\mathbb{R}^n$.

Then $f(x^*) = f(x^*) + \sum_1^m \lambda_i^* g_i(x^*)$ $\quad (\lambda_i^* g_i(x^*) = 0, \forall i)$

$= s(x^*)$

$\leq s(x)$ for any $x \in \mathbb{R}^n$

$= f(x) + \sum_1^m \underset{\overset{\vee\!\shortparallel}{0}}{\lambda_i} g_i(x)$

$\leq f(x)$ if x is feasible for $(*)$, as $g_i(x) \leq 0$. Q.E.D.

If you find this argument interesting or tricky, it is because it is. In fact, there is a tricky connection of the result above to a seemingly unrelated topic called DUALITY.

Second Order Conditions

$\{d : c_i(x)^T d \geq 0,\ i \in \mathcal{I} \cap \mathcal{A}(x^*),\ c_i(x)^T d = 0,\ i \in \mathcal{E}\}$

If $d \in \mathcal{F}(x^*)$ is s.t. $\nabla f(x^*)^T d > 0$, we know $f(x^* + \varepsilon d) > f(x^*)$, small $\varepsilon > 0$.

If $d \in \mathcal{F}(x^*)$ is s.t. $\nabla f(x^*)^T d = 0$, we cannot tell from 1st derivative information alone if $f \uparrow$ or $\downarrow$ when moving from x^* in direction d.

From now on:

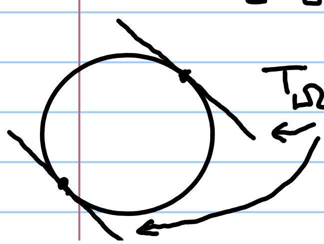
Assume f and c_i, $i \in \mathcal{E} \cup \mathcal{I}$, are C^2, so we can use 2nd derivative information.

Recall our first two examples:

$$\min\ x_1 + x_2 \quad \text{s.t. } x_1^2 + x_2^2 - 1 = 0$$

KKT pts: $x^* = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$, $x^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\nabla f = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -\frac{1}{2}\begin{bmatrix} -2 \\ -2 \end{bmatrix}$ $\quad \nabla f = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 2 \\ 2 \end{bmatrix}$



$T_\Omega(x^*) = \mathcal{F}(x^*) = \{d : \begin{bmatrix} 2 \\ 2 \end{bmatrix}^T d = 0\}$

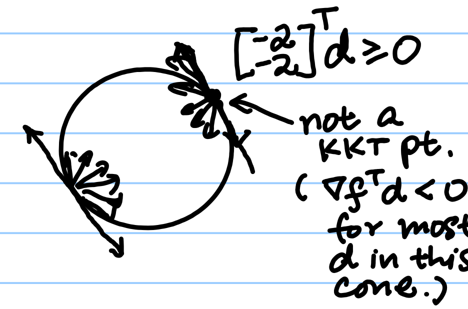
for both x^*.

$$\min\ x_1 + x_2 \quad \text{s.t. } 1 - x_1^2 - x_2^2 \geq 0$$

$x^* = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

$\nabla f = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \underset{\geq 0}{\frac{1}{2}} \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

$T_\Omega(x^*) = \mathcal{F}(x^*) = \{d : \begin{bmatrix} 2 \\ 2 \end{bmatrix}^T d \geq 0\}$



In the first example, KKT cannot distinguish local min $\begin{bmatrix}-1\\-1\end{bmatrix}$ from local max $\begin{bmatrix}1\\1\end{bmatrix}$.

In the second example, KKT can rule out $\begin{bmatrix}1\\1\end{bmatrix}$ (thanks to the inequality).

But in both examples, KKT alone cannot decide if $\begin{bmatrix}1\\1\end{bmatrix}$ is a local minimizer.

At a KKT point x^*, $\nabla f(x^*) = \sum_{i\in\mathcal{E}} \lambda_i^* \nabla c_i(x^*) + \sum_{i\in\mathcal{I}\cap\mathcal{A}(x^*)} \overset{\geq 0}{\lambda_i^*} \nabla c_i(x^*)$

If $d \in \mathcal{F}(x^*)$, $\nabla f(x^*)^T d = \sum_{i\in\mathcal{E}} \lambda_i^* \underbrace{\nabla c_i(x^*)^T d}_{=0} + \sum_{i\in\mathcal{I}\cap\mathcal{A}(x^*)} \lambda_i^* \nabla c_i(x^*)^T d$

If $d\in\mathcal{F}(x^*)$ is such that $\nabla c_i(x^*)^T d > 0$ for some $i\in\mathcal{I}\cap\mathcal{A}(x^*)$ with $\lambda_i^*>0$ then $\nabla f(x^*)^T d > 0$.

no need to worry about such directions

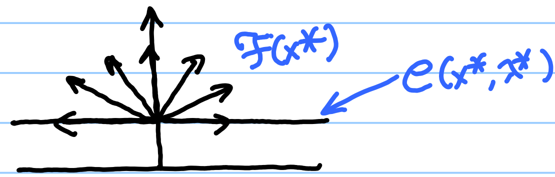
So, the directions we do have to worry about are those in:

$$\boxed{\mathcal{C}(x^*,\lambda^*) := \{ w\in\mathcal{F}(x^*) : \nabla c_i(x^*)^T w = 0,\ i\in\mathcal{A}(x^*)\cap\mathcal{I},\ \lambda_i^*>0\}}$$

$$= \{ w\in\mathbb{R}^n : \nabla c_i(x^*)^T w = 0, \forall i\in\mathcal{E} \text{ and } \nabla c_i(x^*)^T w = 0, \forall i\in\mathcal{I}\cap\mathcal{A}(x^*) \text{ s.t. } \lambda_i^*>0 \text{ and } \nabla c_i(x^*)^T w \geq 0, \forall i\in\mathcal{I}\cap\mathcal{A}(x^*) \text{ s.t. } \lambda_i^*=0\}.$$

E.g. $\min \; x_1^3 + x_2^2$

s.t. $x_2 - 1 \geq 0$



$$\nabla f = \begin{bmatrix} 3x_1^2 \\ 2x_2 \end{bmatrix} \qquad \nabla c_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

KKT conditions : $\begin{bmatrix} 3x_1^2 \\ 2x_2 \end{bmatrix} = \lambda_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Rightarrow x_1 = 0$

a cone, not a subspace

$\lambda_1 (x_2 - 1) = 0$ If $\lambda_1 = 0$, then $x_1 = x_2 = 0$ ← not feasible.

$x_2 \geq 1$ If $\lambda_1 > 0$, then $x_2 = 1$, and $\lambda_1 = 2$.

$x^* = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\lambda_1^* = 2$, $\mathcal{F}(x^*) = \left\{ d : \begin{bmatrix} 0 \\ 1 \end{bmatrix}^T d \geq 0 \right\} = \left\{ \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} : d_2 \geq 0 \right\}$.

the set of directions we need to worry about → $\mathcal{C}(x^*, \lambda^*) = \left\{ \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} : d_2 = 0 \right\}$. ← not only a cone but also a subspace

Unlike our second example, x^* is not a local minimizer.

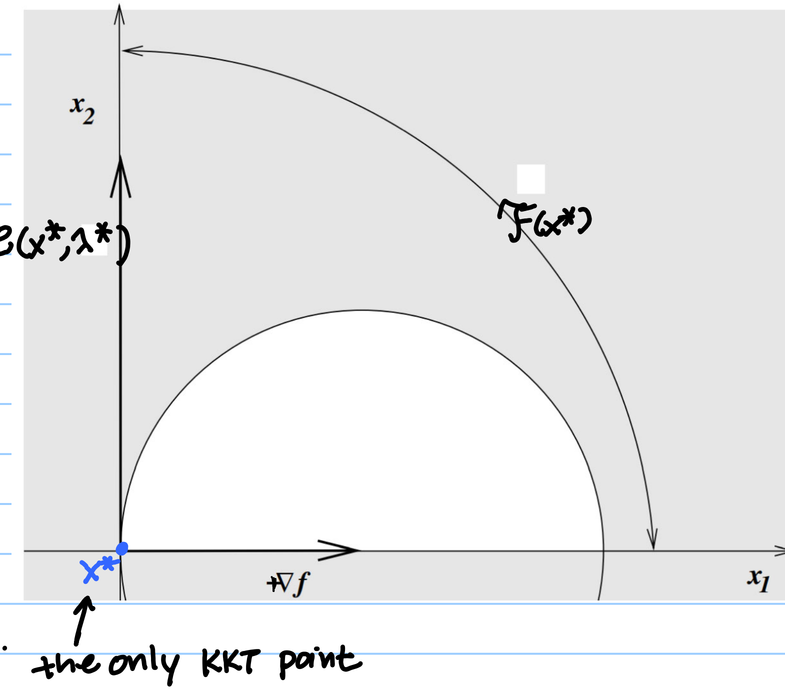
E.g. $\min x_1$ s.t. $x_2 \geq 0, \; 1-(x_1-1)^2 - x_2^2 \geq 0$

$x^* = \begin{bmatrix}0\\0\end{bmatrix}$, $\nabla f = \begin{bmatrix}1\\0\end{bmatrix}$, $\nabla c_1 = \begin{bmatrix}0\\1\end{bmatrix}$, $\nabla c_2(x^*) = \begin{bmatrix}2\\0\end{bmatrix}$.

$\mathcal{F}(x^*) = \{ d : d \geq 0 \}$

$\nabla f(x^*) = 0 \cdot \nabla c_1(x^*) + \frac{1}{2} \nabla c_2(x^*)$, $\lambda^* = \begin{bmatrix}0\\1/2\end{bmatrix}$

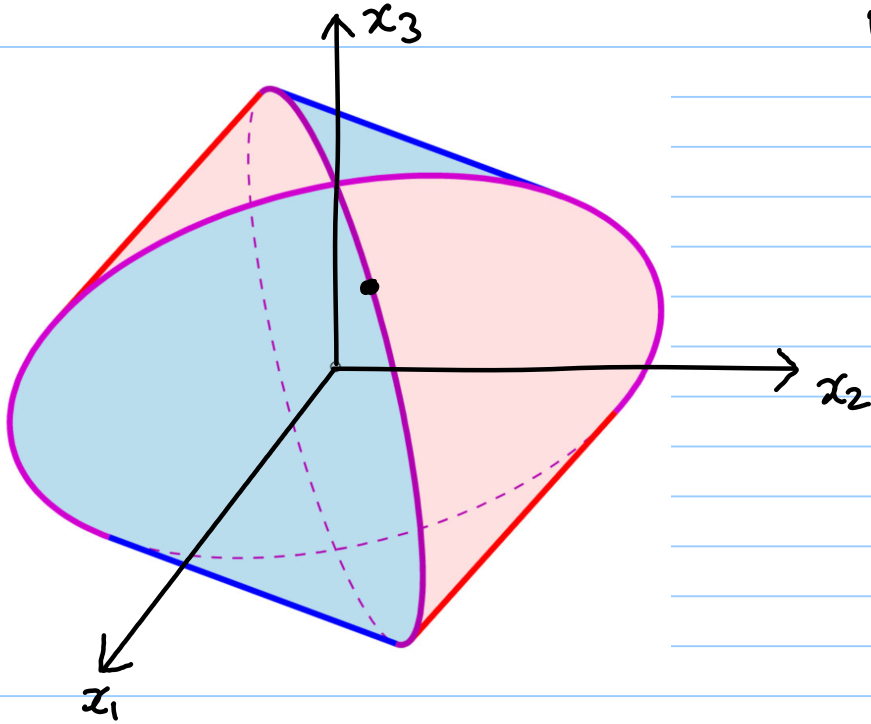
$\mathcal{C}(x^*, \lambda^*) = \left\{ \begin{bmatrix}0\\w_2\end{bmatrix} : w_2 \geq 0 \right\}$. $\leftarrow$ a cone, not a subspace.



This is also an example of non-strict complementarity.

Definition 12.5 (Strict Complementarity).

Given a local solution x^* of (12.1) and a vector λ^* satisfying (12.34), we say that the strict complementarity condition *holds if* exactly one *of λ_i^* and $c_i(x^*)$ is zero for each index $i \in \mathcal{I}$. In other words, we have that $\lambda_i^* > 0$ for each $i \in \mathcal{I} \cap \mathcal{A}(x^*)$.*



HW:

Consider $\max \ x_1 + x_2 + x_3$

s.t. $x_1^2 + x_3^2 \leq 1$

$x_2^2 + x_3^2 \leq 1$

Determine the KKT point(s) x^*.

Determine $\mathcal{F}(x^*)$. Is it a subspace?

Determine $\mathcal{C}(x^*, \lambda^*)$. Is it a subspace?

Is x^* a local maximizer?
(Hint: convexity.)

Theorem 12.5 (Second-Order Necessary Conditions).

Suppose that x^* is a local solution of (12.1) and that the LICQ condition is satisfied. Let λ^* be the Lagrange multiplier vector for which the KKT conditions (12.34) are satisfied. Then

$$w^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) w \geq 0, \qquad \text{for all } w \in \mathcal{C}(x^*, \lambda^*). \tag{12.57}$$

Theorem 12.6 (Second-Order Sufficient Conditions).

Suppose that for some feasible point $x^* \in \mathbb{R}^n$ there is a Lagrange multiplier vector λ^* such that the KKT conditions (12.34) are satisfied. Suppose also that

$$w^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) w > 0, \qquad \text{for all } w \in \mathcal{C}(x^*, \lambda^*),\ w \neq 0. \tag{12.65}$$

Then x^* is a strict local solution for (12.1).

Q: why shouldn't the condition be $w^T \nabla^2 f(x^*) w \geq 0$, $w \in \mathcal{C}(x^*, \lambda^*)$?

A: with constraints, $\nabla f(x^*) \neq 0$ in general.

Key idea of the proofs :

originally, we just view it as a notational trick to write $\nabla f(x^*) = \sum_i \lambda_i^* \nabla c_i(x^*)$

$$\mathcal{L}(x,\lambda) = f(x) - \sum_i \lambda_i c_i(x^*)$$

At a KKT point x^*, λ^*, $\lambda_i^* c_i(x^*) = 0 \ \forall i$, $\underbrace{\nabla_x \mathcal{L}(x^*, \lambda^*) = 0}$

$\mathcal{L}(x^*, \lambda^*) = f(x^*)$,

For $w \in \mathcal{C}(x^*, \lambda^*) \subset \mathcal{F}(x^*)$, from an argument from Step II of the proof of KKT,

$$\exists \ z_k \in \Omega, \ t_k > 0 \ \text{ st. } \lim_{k\to\infty} \frac{z_k - x^*}{t_k} = w \Leftrightarrow z_k - x^* = t_k w + o(t_k)$$

And recall that such z_k is constructed to satisfy the nonlinear equations

$$c_i(z_k) = t_k \nabla c_i(x^*)^T w \quad \forall i \in \mathcal{A}(x^*).$$

For such $z_k \approx x^*$, we have
$$\mathcal{L}(z_k, \lambda^*) = f(z_k) - \sum_{i\in\mathcal{E}\cup\mathcal{I}} \lambda_i^* c_i(z_k)$$
$$= f(z_k) - t_k \sum_{i\in\mathcal{A}(x^*)} \lambda_i^* \nabla c_i(x^*)^T w$$
$$\mathcal{L}(z_k, \lambda^*) = f(z_k)$$

$$\overset{=f(z_R)}{\mathcal{L}(z_R,\lambda^*)} = \overset{=f(x^*)}{\mathcal{L}(x^*,\lambda^*)} + \cancelto{=0}{\nabla_x\mathcal{L}(x^*,\lambda^*)}^T (z_R-x^*) + \frac{1}{2}(z_R-x^*)^T\nabla^2_{xx}\mathcal{L}(x^*,\lambda^*)(z_R-x^*) + o(t_R^2)$$

So

$$f(z_R) = f(x^*) + \frac{1}{2}(z_R-x^*)^T\left[\nabla^2_{xx}\mathcal{L}(x^*,\lambda^*)\right](z_R-x^*) + o(t_R^2) \quad — (*)$$

This explains why the Hessian of $\mathcal{L}(\cdot,\lambda^*)$ is relevant.

In contrast, $f(z_R) = f(x^*) + \underbrace{\nabla f(x^*)^T(z_R-x^*)}_{\neq 0} + \frac{1}{2}(z_R-x^*)^T\nabla^2 f(x^*)(z_R-x_R) + o(t_R^2)$.

Armed with $(*)$, the rest of the proofs is essentially the same as the unconstrained case.

E.g. Again, consider $\min x_1+x_2$ s.t. $2-x_1^2-x_2^2 \geq 0$

KKT point : $x^* = \begin{bmatrix}-1\\-1\end{bmatrix}$, $\lambda_1^* = \frac{1}{2}$

$\mathcal{L}(x,\lambda) = (x_1+x_2) - \lambda_1(2-x_1^2-x_2^2)$, $\nabla_{xx}^2 \mathcal{L}(x^*,\lambda^*) = \begin{bmatrix}2\lambda_1^* & 0\\ 0 & 2\lambda_1^*\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \succ 0.$

(we don't even need to restrict to the directions in $\mathcal{C}(x^*,\lambda^*)$.)

By Thm 12.6, $x^* = \begin{bmatrix}-1\\-1\end{bmatrix}$ is a strict local minimizer.)

[But since the optimization problem is convex, x^* is also a global minimizer.]

Since it is an overkill to apply Thm 12.6 to the convex problem above, let's consider the non-convex problem :

E.g. $\min -\frac{1}{10}(x_1-4)^2 + x_2^2$ s.t. $x_1^2+x_2^2-1 \geq 0$ (neither f nor Ω is convex.)

Note : f is not bounded below on $\{x : x_1^2+x_2^2-1 \geq 1\}$, $f\left(\begin{bmatrix}x_1\\0\end{bmatrix}\right) \to -\infty$, $x_1 \to +\infty$. or $x_1 \to -\infty$

Let's look for local minimizer(s).

$$\mathcal{L}(x,\lambda) = -\tfrac{1}{10}(x_1-4)^2 + x_2^2 - \lambda_1(x_1^2+x_2^2-1)$$

$$\nabla_x \mathcal{L}(x,\lambda) = \begin{bmatrix} -\tfrac{1}{5}(x_1-4) - 2\lambda_1 x_1 \\ 2x_2 - 2\lambda_1 x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \nabla^2_{xx}\mathcal{L}(x,\lambda) = \begin{bmatrix} -\tfrac{1}{5} - 2\lambda_1 & 0 \\ 0 & 2-2\lambda_1 \end{bmatrix}.$$

and $\lambda_1(x_1^2+x_2^2-1)=0$

If $\lambda_1 = 0$, $x_1 = 4$, $x_2 = 0$

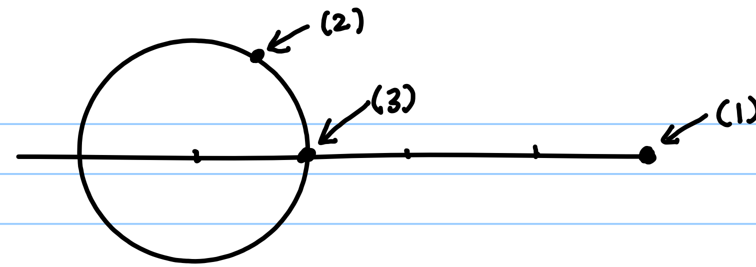
If $\lambda_1 > 0$, $x_1^2+x_2^2=1$, $-\tfrac{1}{5}(x_1-4) - 2\lambda_1 x_1 = 0$ and $(1-\lambda_1)x_2 = 0 \Rightarrow \lambda_1 = 1$ or $x_2 = 0$

$\lambda_1 = 1 \Rightarrow x_1 = 4/11 \Rightarrow x_2 = \frac{\sqrt{105}}{11}$

$x_2 = 0 \Rightarrow x_1 = 1 \Rightarrow \lambda_1 = 3/10$

$$\boxed{\text{KKT points : (1) } x^* = \begin{bmatrix} 4 \\ 0 \end{bmatrix}, \lambda^* = 0, \text{ (2) } x^* = \tfrac{1}{11}\begin{bmatrix} 4 \\ \sqrt{105} \end{bmatrix}, \lambda^* = 1, \text{ (3) } x^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \lambda^* = 3/10.}$$

$\begin{bmatrix} 4 \\ 0 \end{bmatrix}$ is clearly not a local min.



2nd KKT point:

$$\nabla_{xx}^2 \mathcal{L}\left(\frac{1}{11}\begin{bmatrix}4\\ \sqrt{105}\end{bmatrix}, 1\right) = \begin{bmatrix}-2 & 0\\ 0 & 0\end{bmatrix}.$$

$$\nabla c_1(x^*) = 2x^* = \frac{2}{11}\begin{bmatrix}4\\ \sqrt{105}\end{bmatrix}, \quad \mathcal{C}(x^*, \lambda^*) = \{w : 4w_1 + \sqrt{105}\, w_2 = 0\}$$

$$= \left\{ t\begin{bmatrix}\sqrt{105}\\ -4\end{bmatrix} : t \in \mathbb{R}\right\}$$

So $w^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) w = t^2 \begin{bmatrix}\sqrt{105} & -4\end{bmatrix}\begin{bmatrix}-2 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\sqrt{105}\\ 4\end{bmatrix} = -2\cdot(105)\, t^2 < 0$ when $t \neq 0$

By the necessity theorem, the second KKT pt cannot be a local minimizer.

3rd KKT point:

$$\nabla_{xx}^2 \mathcal{L}\left(\begin{bmatrix}1\\ 0\end{bmatrix}, \frac{3}{10}\right) = \begin{bmatrix}-0.4 & 0\\ 0 & 1.4\end{bmatrix}$$

$$\nabla c_1(x^*) = \begin{bmatrix}2\\ 0\end{bmatrix}, \quad \mathcal{C}(x^*, \lambda^*) = \left\{ t\begin{bmatrix}0\\ 1\end{bmatrix} : t \in \mathbb{R}\right\}$$

$$w^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) w = [0, t]\begin{bmatrix}-0.4 & 0\\ 0 & 1.4\end{bmatrix}\begin{bmatrix}0\\ t\end{bmatrix} = 1.4t^2 > 0 \quad \forall\, t \neq 0.$$

By the sufficiency theorem, the third KKT point is a local minimizer.

So we now have a satisfactory generalization of the $f'=0$, $f''>0$ conditions for 1-D unconstrained problems all the way to multi-dimensional optimization problems with equality and/or inequality constraints.

Recall that for unconstrained multivariate problems $f''>0$ is replaced by the condition:

$$\nabla^2 f(x^*) \succ 0 \qquad (\text{i.e. } w^T \nabla^2 f(x^*) w > 0 \ \forall w \in \mathbb{R}^n \setminus \{0\})$$

If we are in (say) $n=100$ dimension, how are we supposed to check this condition?

The spectral theorem tells out that this condition is equivalent to that all the eigenvalues of $\nabla^2 f(x^*)$ are positive.

Moreover, there are well-established techniques, implemented in sophistated software packages, for computing the eigenvalues of symmetric matrices.

Bad news : There is no known polynomial (in n) time algorithm for Checking if a symmetric matrix A satisfies

$$w^T A w \geqslant 0, \quad \forall \underbrace{w \geq 0}_{\text{a cone}}.$$

"Copositivity"

Mathematical Programming 39 (1987) 117–129
North-Holland

SOME NP-COMPLETE PROBLEMS IN QUADRATIC AND NONLINEAR PROGRAMMING

Katta G. MURTY*

Department of Industrial and Operations Engineering, The University of Michigan, 1205 Beal Avenue, Ann Arbor, MI 48109-2117, USA

Santosh N. KABADI**

Faculty of Administration, University of New Brunswick, Fredericton, NB, Canada E3B 5A6

Received 13 December 1985
Revised manuscript received 9 March 1987

In continuous variable, smooth, nonconvex nonlinear programming, we analyze the complexity of checking whether

(a) a given feasible solution is not a local minimum, and

(b) the objective function is not bounded below on the set of feasible solutions.

We construct a special class of indefinite quadratic programs, with simple constraints and integer data, and show that checking (a) or (b) on this class is NP-complete. As a corollary, we show that checking whether a given integer square matrix is not copositive, is NP-complete.

Key words: Nonconvex nonlinear programming, local minimum, global minimum, copositive matrices, NP-complete.

Fortunately, under the LICQ and strict complementarity conditions,

$$\mathcal{C}(x^*, \lambda^*) = \{ w : \nabla c_i(x^*)^T w = 0, \ \forall i \in \mathcal{A}(x^*) \}$$

← not only a cone, but a subspace with $\dim = n - |\mathcal{A}(x^*)|$

$$= \text{null } A(x^*)$$

$$= \{ Zu : u \in \mathbb{R}^{n - |\mathcal{A}(x^*)|} \}$$

Z: columns form a basis of null $A(x^*)$

$$A(x^*) := \begin{array}{c} \begin{bmatrix} \vdots \\ - \nabla c_i(x^*)^T - \\ \vdots \end{bmatrix}_{i \in \mathcal{A}(x^*)} \end{array}$$

rank $|\mathcal{A}(x^*)|$

nullity $n - |\mathcal{A}(x^*)|$

Then

$$w^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) w \geq 0 \quad \forall w \in \mathcal{C}(x^*, \lambda^*)$$

$$\iff u^T Z^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) Z u \geq 0, \ \forall u \in \mathbb{R}^{\dim \mathcal{C}(x^*)}$$

$$\iff Z^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) Z \text{ is positive semidefinite}$$

$$\iff \text{all eigenvalues of } Z^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) Z \text{ are non-negative}$$

When $\mathcal{C}(x^*, \lambda^*)$ is a subspace, standard linear algebra tools can be used to check the 2nd order conditions.