Constrained Optimization (Lecture 5)

March 31st, 20

Last time → Second order: necessary / sufficient > conditions for optimu

→ case of Linear Equality constraints.

using elimination of variables.

(1) $\begin{cases} \min & f(\underline{x}) \\ \text{s.t.} & A\underline{x} = \underline{b} \end{cases}$

A is $m \times n$, $m < n$.

Today

- Example of using elimination of variables for equality constrained problem
- The KKT system for the case when in (1), f is quadratic.
- Using the KKT system for the general case (1).

Example of Equality constrained (solved by elimination of variables).

$$\begin{cases} \min & q(\underline{x}) = x_1^2 - 2x_1 + x_2^2 - x_3^2 + 4x_3 \\ \text{s.t.} & x_1 - x_2 + 2x_3 = 2 \end{cases}$$

Let's eliminate x_1. (solve for x_1, in terms of x_2 & from the constraint, and replace it in the obj. function

$$x_1 = x_2 - 2x_3 + 2.$$

So

$$q(x_2 - 2x_3 + 2, x_2, x_3) = \varphi(x_2, x_3) =$$

$$= 2x_2^2 + 3x_3^2 - 4x_2x_3 + 2x_2$$

→ the reduced objective function

Let's rename the variables:

$x_2 \to \alpha_1$

$x_3 \to \alpha_2$.

So, we want to minimize the <u>unconstrained</u> function:

$$\boxed{\varphi(\underline{\alpha}) = \varphi(\alpha_1, \alpha_2) = 2\alpha_1^2 + 3\alpha_2^2 - 4\alpha_1\alpha_2 + 2\alpha_1}$$

Set $\nabla\varphi(\underline{\alpha}) = \underline{0}$ and solve for the stationary points.

(first-order necessary cond. for unconstr. optimal.

But from last time, we showed that:

$$\nabla\varphi(\underline{\alpha}) = Z^T \nabla q(\underline{\hat{x}} + Z\underline{\alpha}) = Z^T \nabla q(\tilde{x})$$

where $\hat{x}$ is a feasible point.

$(\hat{x}_1 = \hat{x}_2 - 2\hat{x}_3 + 2)$.

and $Z = [\underline{z}_1, \underline{z}_2]$ is an 3×2 matrix

original # of des. variables

number of reduced design variables

whose columns are vectors in a basis of Null (A)

Let's find the null (A): set of vectors in $\mathbb{R}^n$ ($\mathbb{R}^3$) that are mapped by A into 0.

$\underset{1\times 3}{A} = [1 \quad -1 \quad 2]$, $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

$A\underline{x} = 0$

↳ is a vector, in general, of dimension equal to the # of equality constraints.

March 31^{st}, 2020

$x_1 - x_2 + 2x_3 = 0.$

solve for x_1, and let x_2 & x_3 be arbitrary:

$$x_1 = x_2 - 2x_3$$

So the set of vectors mapped by A into ~~zero~~ is of this form:

$$\begin{bmatrix} x_2 - 2x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}, \quad x_2, x_3 \in \mathbb{R} \text{ arbitrary.}$$

write it as a linear combination of two basis vectors.

Vectors $\underline{z}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ & $\underline{z}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$ form a basis for Null(A).

So $\boxed{Z = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}}$ $\boxed{Z^T = \begin{bmatrix} 1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}}$

Let's solve $\nabla \varphi(\underline{\alpha}) = \underline{0}.$

$$\nabla \varphi(\underline{\alpha}) = \begin{bmatrix} 4\alpha_1 - 4\alpha_2 + 2 \\ 6\alpha_2 - 4\alpha_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{cases} 4\alpha_1 - 4\alpha_2 = -2 \\ -2\alpha_1 + 3\alpha_2 = 0 \quad | \cdot 2 \end{cases}$$

add together $\rightarrow$ $2\alpha_2 = -2$ $\Rightarrow$ $\boxed{\alpha_2 = -1}$

$\boxed{\alpha_1 = -\frac{3}{2}}$

March 31^{st}, 2020

Going back to the notation, we get :

$x_2 = -\frac{3}{2}$

$x_3 = -1.$

Use the constraint, get .

$$x_1 = x_2 - 2x_3 + 2 = \frac{5}{2}$$

The only stationary point is: $\underline{x}^* = \begin{bmatrix} \frac{5}{2} \\ -\frac{3}{2} \\ -1 \end{bmatrix}$

To see if it is a minimizer, use second-order conditions.

($\nabla^2_{\alpha\alpha} \varphi(\underline{\alpha})$ is P.D. at $\underline{x}^*$)

$\boxed{\text{Exercise}}$: Compute $\nabla^2_{\alpha\alpha} \varphi(\underline{\alpha})$ an see if it is p.s.d, or p.d.

<u>Note</u>: The example we tried here was special.

The objective function was quadratic. (original or the transformed one).

$\Longrightarrow$ $\nabla \varphi(\underline{\alpha}) = \underline{0}$ was a <u>linear system</u>.

(can be solved with any linear solver)

When obj. func. is not quadratic (or linear) $\Rightarrow$ $\boxed{\nabla \varphi(\underline{\alpha}) = \underline{0} \text{ is a nonlinear system.}}$

March 31st, 2020

Q What if instead of <u>elimination of variables</u>, we use the KKT conditions?

Consider the same problem (quadratic f, m linear equality constraints)

$$(2)\begin{cases} \min \frac{1}{2}\underline{x}\cdot G\underline{x} + \underline{x}\cdot\underline{d} \\ \text{s.t. } A\underline{x} = \underline{b} \end{cases}$$

(m equality constraints, A is $m\times n$)

Note: G: does it need to be p.d. for a minimizer to exist?

$\underline{w}\cdot\nabla\mathcal{L}(x^*,\lambda^*)\,\underline{w} \geq 0$ for all $\underline{w}\in C(\underline{x}^*,\underline{\lambda}^*)$

Let $\underline{x}^*$ be a local min for (2)

$\underline{w}$ s.t. $\nabla c_i(\underline{x}^*)\cdot\underline{w} = 0$, $i\in E$

$c_i(\underline{x}) = \underline{a}_i\cdot\underline{x} - b_i$, $i=1,2,\dots m$.

$\underline{a}_i$ is the i^{th} row in matrix A

$\nabla c_i(\underline{x}) = \underline{a}_i$

Want $\underline{w}$, s.t. $\underline{a}_i\cdot\underline{w} = 0$, $i=1,\dots m$

$\Rightarrow$ $A\underline{w} = \underline{0}$, $\underline{w}\in Null(A)$.

KKT conditions

(assume A has full row rank.

$\Rightarrow$ l.i. rows

$\Rightarrow$ LICQ holds)

there exists $\underline{\lambda}^*\in\mathbb{R}^m$ s.t.

$\nabla_{\underline{x}}\mathcal{L}(\underline{x}^*,\underline{\lambda}^*) = \underline{0}$, where $\mathcal{L}(\underline{x},\underline{\lambda}) = \frac{1}{2}\underline{x}\cdot G\underline{x} + \underline{x}\cdot\underline{d} - \underline{\lambda}\cdot(A\underline{x}-\underline{b})$

Let's compute $\nabla_{\underline{x}}\mathcal{L}(\underline{x},\underline{\lambda}) = G\underline{x} + \underline{d} - A^T\underline{\lambda}$.

(indeed: $\nabla_{\underline{x}}(\underline{\lambda}\cdot A\underline{x}) = \nabla_{\underline{x}}(\underbrace{A^T\underline{\lambda}}\cdot\underline{x}) = A^T\underline{\lambda}$.

March 31st, 2020

So, KKT condition say that:

$$\begin{cases} G\underline{x}^* + \underline{d} - A^T\underline{\lambda}^* = \underline{0} \\ A\underline{x}^* = \underline{b} \end{cases} \iff \begin{cases} G\underline{x}^* - A^T\underline{\lambda}^* = -\underline{d} \\ A\underline{x}^* \qquad\quad = \underline{b} \end{cases}$$

$$\begin{array}{c|c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{bmatrix} G & -A^T \\ A & 0 \end{bmatrix} \end{array} \begin{bmatrix} \underline{x}^* \\ \underline{\lambda}^* \end{bmatrix} = \begin{bmatrix} -\underline{d} \\ \underline{b} \end{bmatrix} \begin{array}{c} \}\, n \\ \}\, m \end{array} \qquad (3)$$

$(m+n)$ by $(m+n)$ system of equations. $\boxed{\text{Matrix is } \underline{\text{not}} \text{ symmetric}}$

Q How do we use this in practice? (we don't have $\underline{x}^*$)

A Start at some arbitrary $\underline{x}$.

We want to get to $\underline{x}^*$ from $\underline{x}$.

$$\underline{x} \xrightarrow{\underline{p}} \underline{x}^*$$

So set $\underline{x}^* = \underset{\text{known}}{\underline{x}} + \underset{\text{unknown.}}{\underline{p}}$

Write system (3) and separate ~~unknowns~~ unknowns

First block of eqs.

$$G(\underline{x}+\underline{p}) - A^T \underline{\lambda}^* = -\underline{d}$$

$$G\underline{p} - A^T\underline{\lambda}^* = -\underline{d} - G\underline{x}$$

$$\boxed{G(-\underline{p}) + A^T\underline{\lambda}^* = G\underline{x} + \underline{d}}$$

Second block of m eqs.

March 31^{st}, 2020

$$A\underline{x} + A\underline{p} = \underline{b} \quad \Rightarrow \quad A\underline{p} + 0\,\underline{\lambda}^* = \underline{b} - A\underline{x}$$

$$A(-\underline{p}) + 0\,\underline{\lambda}^* = A\underline{x} - \underline{b}.$$

System becomes:

$$\boxed{\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} -\underline{p} \\ \underline{\lambda}^* \end{bmatrix} = \begin{bmatrix} \underline{g} \\ \underline{c} \end{bmatrix}},$$

with

$\underline{g} = G\underline{x} + \underline{d}$

$\underline{c} = A\underline{x} - \underline{b}$.

Remark: This system is symmetric! (more efficient to solve than non-sym. ones)

Note: • We will not be able to use Cholesky or C.G. to solve. Why?

(Matrix is sym. but not p.d.) zeroes on diagonal.

even if G is p.d. ↳ (not necessary).

• Can use GEPP, or specialized methods.

does not make use of symmetry.

↳ - Null-space Method
- Schur-complement method.

or iterative methods: ⇒ projected C.G.

Sections 16.2, 16.3 in textbook

March 31^{st}, 2020

$\boxed{\text{Note}}$ what if f is nonlinear in general?

(4) $\begin{cases} \min f(\underline{x}) \\ \text{s.t. } A\underline{x} = \underline{b} \end{cases}$

Can start with some $(\underline{x}_0, \underline{\lambda}_0)$ then, at iteration $\underline{k}$ (we know $\underline{x}_k$), solve

$$\begin{cases} \min \ \frac{1}{2}\, \underline{x} \cdot \nabla^2 f(\underline{x}_k)\, \underline{x} + \nabla f(\underline{x}_k) \cdot \underline{x} + f(\underline{x}_k) \\ \text{s.t. } A\underline{x} = \underline{b} \end{cases}$$

KKT system is:

$$\begin{bmatrix} \nabla^2 f(\underline{x}_k) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} -\underline{p}^* \\ \underline{\lambda}^* \end{bmatrix} = \begin{bmatrix} G\underline{x}_k + \underline{d} \\ A\underline{x}_k - \underline{b} \end{bmatrix}$$

then $\underline{x}_{k+1} = \underline{x}_k + \underline{p}$

$\underline{\lambda}_{k+1} = \underline{\lambda}_k + \underline{\lambda}$

Remark

only $Z^T \nabla^2 f(\underline{x}_k^*) Z$ needs to be p.s.d, not the entire $\nabla^2 f(\underline{x}_k)$.

where $Z \rightarrow$ matrix whose columns form a basis for Null(A)

$Z = [\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_p]$

$\underline{w} = \alpha_1 \underline{z}_1 + \cdots + \alpha_p \underline{z}_p$

$p = n - m$

March 31st, 2020

Remark We now know how to solve problem with <u>Linear</u> <u>Equality</u> Constraints.

(Q) Can we use elimination of variables method for <u>Nonlinear</u> <u>Equality</u> Constraints?

(A) Not always; It's not safe, <u>So we will not use it</u>.

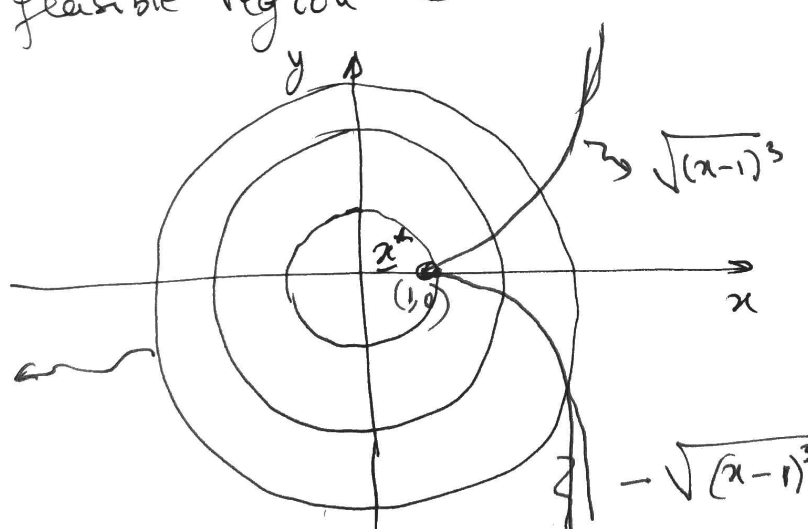
<u>Example</u>:

$$(5)\quad \begin{cases} \min \ x^2+y^2 \\ \text{s.t.} \\ \quad (x-1)^3 = y^2 \end{cases}$$

Graphical method first:
Draw feasible region & level curves:

$y^2 = (x-1)^3 \qquad \underline{x \geq 1}$

$y = \pm\sqrt{(x-1)^3}$



level curves for x^2+y^2 ($x^2+y^2 = \text{constant}$)

<u>exact solution</u>: $x^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Let's try to solve it using elimination of variables:

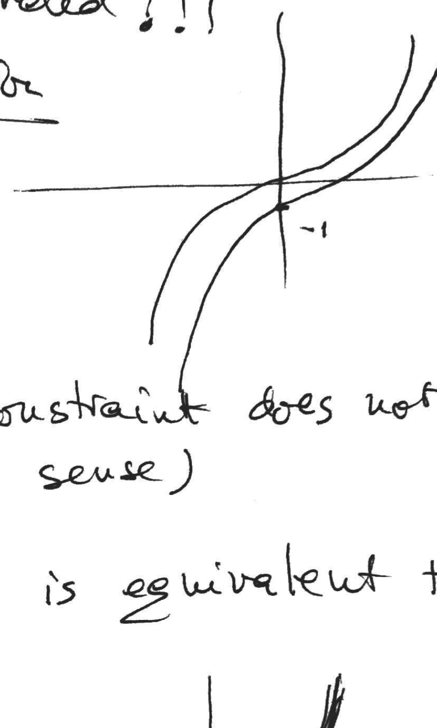
March 31st, 2020

$$\min \quad x^2 + \underset{\uparrow \atop (x-1)^3}{y^2} = x^2 + (x-1)^3 = x^3 - 3x^2 + 3x - 1 + x^2 =$$

$$= x^3 - 2x^2 + 3x - 1$$

unbounded !!!

<u>Not solution</u>

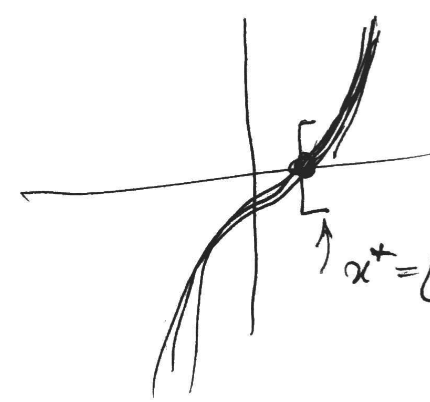


What happened ?!?

We lost the (implicit) condition that <u>$x \geq 1$</u> (otherwise the constraint does not make sense)

So, original problem (5) is <u>equivalent</u> to :

$$\begin{cases} \min & x^2 + (x-1)^3 \\ \text{s.t.} & x - 1 \geq 0 \end{cases}.$$



$x^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

<u>Remark</u>: This type of mistake is hard to catch if we have many constraints and n is large ; To avoid danger

$$\boxed{\text{We will not use elimination of variables for nonlinear equality constraints.}}$$

March 31st, 2020

Summary :- showed how to solve minimization problems with ~~eq~~ linear equality constraints.

- Next time: Method for solving

$$\begin{cases} \min f(\underline{x}) & \longrightarrow \text{quadratic} \\ \text{s.t. } c_i(\underline{x}) = 0 \quad i \in E & \longrightarrow \text{Linear} \\ c_i(\underline{x}) \geq 0 \quad i \in I & \longrightarrow \text{Linear} \end{cases}$$

added

→ a "Quadratic programming problem".