§4.1 Optimization Problems

"NLP" form ("Nonlinear Programming")

$$\min_{x \in \mathbb{R}^n} f_0(x) \quad \text{such that ("s.t.")} \quad f_i(x) \le 0 \quad i = 1, \dots, m$$
$$h_i(x) = 0 \quad i = 1, \dots, p$$

Optimal value denoted p^*

($p^* = +\infty$ means infeasible)

= not assuming convexity yet =

If $f_0(x) =$ constant, this is a "feasibility problem"

Sometimes you have to work a bit to put a problem in a standard form (eg. $\max f_0 \iff -\min -f_0$

or $3x \geq 5 \iff 5 - 3x \le 0$)

Equivalent problems means problems w/ same argmin, but otherwise may be different

ex

$$\min \|Ax-b\|_2 \quad \text{vs} \quad \min \|Ax-b\|^2$$

not even differentiable

smooth (strongly smooth!)

Tricks

change of variables (often affine)

Ex:
$$\min_{x \in \mathbb{R}^1_{++}} \varphi(x) \quad \overset{y = 1/x}{\iff} \quad \min_{y \in \mathbb{R}^1_{++}} \tilde{\varphi}(y) := \varphi(1/y)$$

Ex:
$$\min_{x \in \mathbb{R}^1_{+}} \varphi(x) \quad \iff \quad y = \sqrt{x} \quad \min_{y \in \mathbb{R}} \tilde{\varphi}(y) := \varphi(y^2)$$

unconstrained!

Slack variables

$$\min_{x \in \mathbb{R}^n} \; f_0(x) \quad \text{s.t. } f_i(x) \le 0 \quad i=1,\dots,m$$

"s" is a "slack var."

$$\iff \min_{x \in \mathbb{R}^n,\, s \in \mathbb{R}^m} \; f_0(x) \quad \text{s.t.} \quad \begin{array}{l} f_i(x) + s_i = 0, \quad i=1,\dots,m \\ s_i \ge 0, \quad i=1,\dots,m \end{array}$$

Eliminate Constraints

Ex suppose $Ax = b$ is a constraint

$\Rightarrow$ $x = x_0 + y$ for some $y \in \mathrm{Ker}(A)$ $(Ay = 0)$
and a "particular" solution x_0
($Ax_0 = b$. Any such x_0 works)

let F be a matrix whose columns are a basis for $\ker(A)$, then $v = F \cdot z$ for some z

(i.e. if $\underset{\text{SVD}}{A} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix}$, then $F = V_2^T$)

0 sing. values

then $\min_x f_0(x)$ st $Ax = b$ $\iff$ $\min_z f_0(x_0 + Fz)$

(unconstrained!)

Epigraph trick

Ex
$$\min_{x \in C} f_0(x) \iff \min_{x \in C,\, t \in \mathbb{R}} t \quad \text{s.t. } f_0(x) \le t$$

Ex
$$\min_{x \in \mathbb{R}^n} \|Ax-b\|_1 \quad \left(:= \sum_{i=1}^{m} |a_i^T x - b_i| \right)$$

$$\iff$$

$$\min_{x \in \mathbb{R}^n,\, t \in \mathbb{R}^m} \sum_{i=1}^{m} t_i \quad \Big\} \text{ aka } \mathbb{1}^T \cdot t \text{ where } \mathbb{1} \text{ is ones vector}$$

$$\text{s.t. } |a_i^T x - b_i| \le t_i \quad \forall i = 1, \dots, m$$

Misc. Tricks

$$\min_{x} f(x) + g(x) \iff \min_{x,z} f(x) + g(z) \quad \text{s.t. } x = z$$

marginalization aka variable projection

$$\min_{x,z} f(x,z) \iff \min_{x} \Bigg[\underbrace{\min_{z} f(x,z)}_{g(x)} \Bigg] = \min_{x} g(x)$$

§4.2 Convex optimization problems

NLP form

$\min f_0(x)$ ← require f_0 to be a convex function

$f_i(x) \le 0$ ← require f_i to be a convex function

$h_i(x) = 0$

⚠ $f_i(x) \ge 0$ (f_i convex) is not a convex set.

i.e. $h_i(x) \le 0$ AND $h_i(x) \ge 0$

so need h_i and $-h_i$ convex

i.e. h_i must be affine, notation: $a_i^T x = b_i$

Recall... (w/ non-empty, closed convex constraints)

f_0 convex $\Rightarrow$ all local sol'n also global

f_0 strictly cvx $\Rightarrow$ global sol'n is unique

f_0 coercive $\Rightarrow$ global sol'n exists

Optimality conditions

Let $C = \{x : f_i(x) \le 0 \ \forall i = 1, \dots, m \text{ and } a_i^T x \le b_i \ i = 1, \dots, p\}$

$\min f(x)$ ← assume cvx fcn

$x \in C$ ← assume cvx set

By convexity, $\exists$ a subgradient at x $d \in \partial f(x)$

$\forall y \quad f(y) \ge f(x) + \langle d, y - x \rangle$

so $\boxed{\text{Thm}}$ $x \in C$ is a global minimizer iff

$\exists\, d \in \partial f(x)$ s.t. $\forall y \in C, \ \langle d, y - x \rangle \ge 0$

"Euler inequality" (type of VI)

proof of Thm.

$\Longleftarrow$ follows directly by definition of subgradient

$\Longrightarrow$ (for simplicity, we'll assume f is differentiable at x, so $d = \nabla f(x)$)

Suppose $\exists\, y \in C$ s.t. $\langle \nabla f(x), y - x \rangle < 0$

Let $\varphi(t) := f(x + t\cdot(y-x))$ so

$$\varphi'(0) = \langle \nabla f(x), y - x \rangle \quad \text{via chain rule}$$
$$< 0 \quad \text{by assumption.}$$

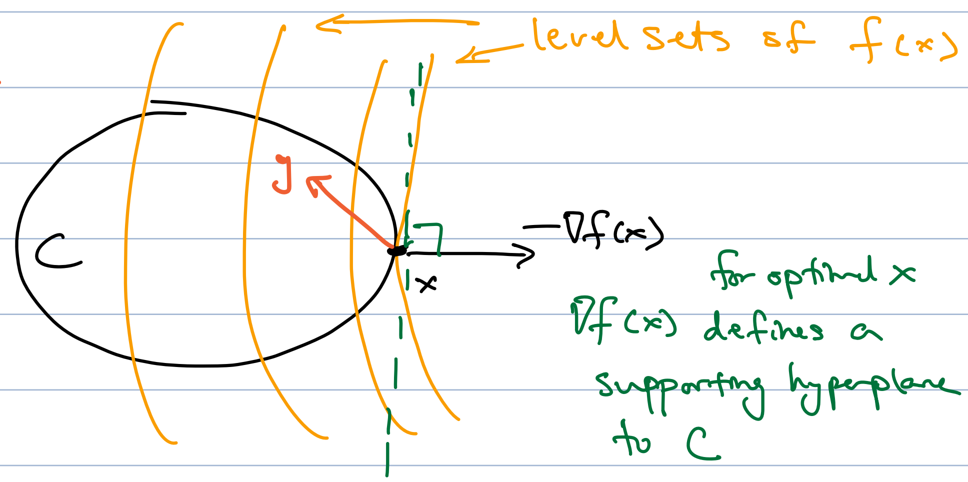
But
$$\varphi'(0) = \lim_{t \to 0^+} \frac{\varphi(t) - \varphi(0)}{t} < 0$$

$$\Rightarrow \exists\, t > 0 \text{ s.t. } \frac{\varphi(t) - \varphi(0)}{t} < 0$$

$$\Rightarrow \varphi(t) < \varphi(0)$$

$\Rightarrow f(x)$ not globally optimal. $\square$

visualization



Optional material: VARIATIONAL INEQUALITIES

ref: "Finite dim. var. ineq. and complementarity problems", Facchinei and Pang, '02

Def let $C \subseteq \mathbb{R}^n$ be closed (often assumed to be convex) and $F: C \to \mathbb{R}^n$ a continuous operator, then the corresponding variational inequality problem is

Find $x \in C$ s.t. $\forall y \in C$, $\langle F(x), y-x \rangle \geq 0$

This generalizes optimization

Recall the normal cone $N_C(x) = \{d \in \mathbb{R}^n : \langle d, y-x \rangle \leq 0 \ \forall y \in C\}$

then VI is looking for $-F(x) \in N_C(x)$

i.e.

$0 \in F(x) + N_C(x)$ so think of F as the subgradient operator

If C is a cone, the VI can be reformulated as a "complementarity problem" (CP):

Find $x \in C$ s.t. $x \perp F(x)$ and $F(x) \in C^*$

($C^* = \{d : \langle d, x \rangle \geq 0 \ \forall x \in C\}$ is dual cone)

ex: $C = C^* = \mathbb{R}^n_+$

and F affine is the Linear Complementarity Problem LCP

Find $x \geq 0$, $\underbrace{Ax+b}_{F(x)} \geq 0$ s.t. $x \perp Ax+b$

$A \in \mathbb{R}^{n \times n}$

$\Rightarrow \forall i$, $x_i = 0$ or $(F(x))_i = 0$ $\Leftarrow$ "complementary"

VI are a non-trivial generalization of optimization

(also, subtle differences: VI look for equivalents of "stationary pts", so if not convex, these aren't equivalent to global minimizers)

eg VI are the KKT equations of optimization problems

ex $\min f(x) := \langle c, x\rangle + \frac{1}{2}\langle x, Qx\rangle$

s.t. $x \geq 0$ Let $Q \succeq 0$ so it's convex

optimality conditions: let $\mathcal{L}(x, y) = f(x) + \langle x, y\rangle$ (y dual variable)

KKT:

$$0 = \nabla_x \mathcal{L}(x, y) = c + Qx + y, \text{ i.e., } y = -Qx - c$$

$$x^T y = 0, \quad x \geq 0, \quad y \geq 0$$

and a generic LCP is solve $\boxed{x \geq 0, \; Ax + b \geq 0, \; x \perp Ax + b}$

$$\iff x \geq 0, \quad \underbrace{Ax+b}_{y} \geq 0,$$

$$x^T \underbrace{(Ax+b)}_{y} = 0$$

... so it's a LCP $A = -Q$, $b = -c$

... but a generic LCP need not have $A \succeq 0$ or $A \preceq 0$: A indefinite is ok, in fact, A need not even be symmetric!

Other ex. of VI include finding Nash Equilibria