den 12 februari 2021 07:51

PSS3 - Optimization

Today: 11.11, exercise; 4.1; 4.5

Lectures Refresh:

$$\min_x f(x) \qquad\qquad (P)$$
$$\text{s.t.} \quad g(x) \preceq 0 \quad \longleftrightarrow \mu$$
$$h(x) = 0 \quad \longleftrightarrow \lambda$$

in general, not possible to solve directly.
instead KKT conditions:

(i) • $\nabla_x L(x^*, \lambda^*, \mu^*) = 0$

(ii) • $g(x^*) \preceq 0$

(iii) • $h(x^*) = 0$

(iv) • $\mu_i^* g_i(x^*) = 0, \quad i = 1, \dots, m$

(v) • $\mu_i^* \geq 0$

to check if (x^*, λ^*, μ^*) is a solution of P.

BUT KKT conditions are just necessary conditions!

There are some exceptions:

- convex problems (i.e. $f(x)$ convex; $g(x)$ convex; $h(x)$ affine) where KKT conditions $\Leftrightarrow$ global optimality.

- QP $\in$ convex problems.

II.II.1 Exercise:

Show that: $\min_x \|Ax\|_1$ can be written as the LP:

$x \in \mathbb{R}^n$; A $m \times n$.

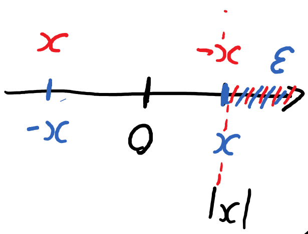
$$\min \; c^T z \quad \text{s.t.} \; Fz \leq g$$

remember that $\|x\|_1 = \sum_{i=1}^{n} |x_i|$.

remark: for any $x \in \mathbb{R}$, $|x|$ is the smallest ε s.t.

$$x \leq \varepsilon \text{ and } -x \leq \varepsilon$$

build some intuition:



$x \in \mathbb{R}$

$$|x| = \min_{\varepsilon} \varepsilon \quad \text{s.t.} \quad x \leq \varepsilon, \; -x \leq \varepsilon$$

$$\min |x| = \min_{\varepsilon, x} \varepsilon \quad \text{s.t.} \quad x \leq \varepsilon, \; -x \leq \varepsilon$$

can be generalized component by component:

$$\min_{x,\varepsilon} \mathbb{1}_m^T \varepsilon \quad \text{with } \varepsilon \in \mathbb{R}^m$$

$$\text{s.t.} \begin{bmatrix} A & -I \\ -A & -I \end{bmatrix} \begin{bmatrix} x \\ \varepsilon \end{bmatrix} \leq 0$$

SCRIBLINGS DON'T READ THIS

$$\min_x \|Ax\|_1 = \sum_{i=1}^{m} |\underbrace{A_i x}_{=y_i}| = \sum_{i=1}^{m} |y_i| = \mathbb{1}^T \varepsilon$$

$$[1 \; 1 \; 1 \ldots 1] \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_m \end{bmatrix}$$

$$\min_x \|Ax\|_1 = \min \sum_{i=1}^{m} |y_i| = \min \sum_{i=1}^{m} \varepsilon_i \quad \text{s.t.} \; y_i \leq \varepsilon_i, \; -y_i \leq \varepsilon_i, \; i = 1, \ldots, m$$

so we can identify:

$$z = \begin{bmatrix} x \\ \varepsilon \end{bmatrix}; \quad c = [0_n \;\; \mathbb{1}_m]^T$$

$$F = \begin{bmatrix} A & -I \\ -A & -I \end{bmatrix}; \quad g = 0_{2m}$$

$$\begin{bmatrix} -A & -I \end{bmatrix}\begin{bmatrix} \varepsilon \end{bmatrix} \leq 0$$

$$F = \begin{bmatrix} A & -I \\ -A & -I \end{bmatrix} ; \quad g = 0_{2m}$$

$\mathbb{1}_m$: vector of ones

0_m : vector of zeros

EXERCISE 4.1

$$\min V(x) = x_1^2 - x_1 x_2 + x_2^2 - 3x_1 \qquad (P\text{-}I)$$

$$\text{s.t.} \quad x_1 \geq 0$$

$$x_2 \geq 0$$

$$x_1 + x_2 \leq 2$$

Standard QP form is: $\min V(x) = \frac{1}{2} x^T Q x + p^T x$

$$\text{s.t.} \quad Gx \leq h$$

$$Ax = b$$

a) Rewrite (P-I) in standard form.

$$V(x) = \frac{1}{2} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \underbrace{\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}}_{Q} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \underbrace{\begin{bmatrix} -3 \\ 0 \end{bmatrix}^T}_{p^T} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Q must be symmetric

There are 3 contraints: $\underbrace{\begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 1 \end{bmatrix}}_{G} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \underbrace{\begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}}_{h}$

$A = 0, \; b = 0$

We have identified $Q, p, G, h, A,$ and b.

b) Show that (P-I) is convex.

$(f(x) \text{ convex} \Leftrightarrow \nabla_x^2 f(x) \succeq 0)$ convexity criteri[on]

($f(x)$ convex $\Leftrightarrow \nabla_x^2 f(x) \succcurlyeq 0$) convexity criterion

for a QP: $\nabla_x^2 V(x) = Q$ so we have to check that $Q \succcurlyeq 0$.

Here: $\mathrm{eig}(Q) = \{1; 3\}$ so Q is positive-definite, and V is convex.

And linear inequalities $\Rightarrow$ convex feasible set.

convex objective function + convex feasible set $\Rightarrow$ convex problem.

c) $x_1^o = 1.5$; $x_2^o = 0.5$.

d) KKT conditions:

μ^o unknown at this point.

(iii) no equality constraints. ok

(ii) does $Gx^o \leqslant h$ $\quad Gx^o - h = \begin{bmatrix} -1.5 \\ -0.5 \\ 0 \end{bmatrix} \leqslant 0$ ok

(iv) $\begin{cases} -1.5\mu_1^o = 0 \\ -0.5\mu_2^o = 0 \\ \mu_3^o \times (\underbrace{x_1^o + x_2^o}_{=2} - 2) = 0 \end{cases} \Rightarrow \mu_1^o = \mu_2^o = 0$ ok

ok

(v) $\mu_3^o \geq 0$

(i) $\nabla_x \mathcal{L} = Qx^o + p + G^T\mu^o = 0$

i.e. $\begin{bmatrix} -0.5 \\ -0.5 \end{bmatrix} + \begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \mu_3^o \end{bmatrix} = 0$ ok

so: $\mu_3^o = 0.5$.

$\mu^o = \begin{bmatrix} 0 \\ 0 \\ 0.5 \end{bmatrix}$; $x^o = \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix}$ is the (global) minimizer.

Remark: $\mu_3^o > 0$ i.e. 3rd constraint is active.

$\mu_1^* = \mu_2^* = 0$ so 1^{st} and 2^{nd} constraints not active.

$\boxed{4.5}$ $\min\ V(x) = \frac{1}{2}x^TQx + p^Tx, \quad Q \succ 0$

s.t. $Ax = b$

Lagrange function: $L(x,\lambda) = \underbrace{\frac{1}{2}x^TQx + p^Tx}_{V(x)} + \underbrace{\lambda^T}_{r}\underbrace{(Ax-b)}_{h(x)}$

Skip a) and b)

c) Form and solve the dual problem.

$$q(\lambda) = \min_x L(x,\lambda)$$

Lagrange dual function: $q(\lambda)$

$L(x,\lambda)$ is convex w.r.t. x (assume λ fixed)

minimize through $\frac{\partial L}{\partial x} = 0$.

$$\nabla_x L(x^*,\lambda) = 0 \iff Qx^* + p + A^T\lambda = 0 \qquad Q \succ 0$$

$$\iff x^* = -Q^{-1}(p + A^T\lambda)$$

so: $q(\lambda) = \frac{1}{2}(p + A^T\lambda)Q^{-1}(p+A^T\lambda) - p^TQ^{-1}(p+A^T\lambda) - \lambda^T(AQ^{-1}(p+A^T\lambda) + b)$

$$= -\frac{1}{2}(p + A^T\lambda)^TQ^{-1}(p+A^T\lambda) - \lambda^Tb$$

Dual problem: $\max_\lambda\ q(\lambda) \iff \min_\lambda\ -q(\lambda)$

no constraints (because only equality constraints in our primal problem).

in general, $q(\lambda)$ is concave.

so dual PB is a concave unconstrained problem: solve it through $\nabla_\lambda g(\lambda^*) = 0$

λ^* $\Rightarrow$ compute x^* dual solution (x^*, λ^*).

One could check that strong duality holds in the end.