KKT equations & complementary slackness

Friday, March 5, 2021 10:15 AM

§5.5 in Boyd + Vandenberghe

Complementary Slackness

Suppose we have primal + dual optimal solutions x^*, λ^*, ν^*, and have strong duality
(no need for convexity) (ie., ∃ saddle pts) $\lambda^* \geq 0$

Observe

$$p^* := \min_{\substack{x \\ f_i(x) \leq 0 \\ h_i(x) = 0}} f_0(x) = f_0(x^*) = g(\lambda^*, \nu^*) \text{ via strong duality}$$

$$= \inf_x \mathcal{L}(x, \lambda^*, \nu^*) \quad \text{by def'n of } g$$

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

x^* is feasible so $h_i(x^*) = 0$

$$= f_0(x^*) + \sum \underbrace{\overbrace{\lambda_i^*}^{\geq 0} \overbrace{f_i(x^*)}^{\leq 0}}_{\leq 0} + \sum \cancel{\nu_i^* h_i(x^*)}^{=0}$$

must "="

$$= (\leq)\ f_0(x^*) = p^*$$

Conclusion: $\forall i$, $\lambda_i^* \cdot f_i(x^*) = 0$ "complementary slackness"

i.e., either $\lambda_i^* = 0$ or $f_i(x^*) = 0$ (or both)

in particular, if $f_i(x^*) < 0$ ("not tight")

$\Rightarrow \lambda_i^* = 0$

also, $x^* \in \operatorname{argmin}_x \mathcal{L}(x, \lambda^*, \nu^*)$

KKT conditions

Karush-Kuhn-Tucker

KKT Thm 1 "necessary" (No convexity needed). [Assume f_i, h_i are differentiable]

If x^* is primal optimal, and λ^*, ν^* dual optimal,
and no duality gap (i.e., strong duality) (i.e., ∃ saddle pts and strong duality)
then necessarily x^*, λ^*, ν^* satisfy

① "stationarity" $0 = \nabla f_0(x^*) + \sum \lambda_i^* \nabla f_i(x^*) + \sum \nu_i^* \nabla h_i(x^*)$

i.e., $0 = \nabla_x \mathcal{L}(x^*, \lambda^*, \nu^*)$

② "primal feasibility" $f_i(x^*) \leq 0$, $h_i(x^*) = 0$

③ "dual feasibility" $\lambda^* \geq 0$ (i.e., $\lambda_i^* \geq 0 \ \forall i$)

④ "complementary slackness" $\lambda_i^* f_i(x^*) = 0$

Interpretation Ⓐ These are only "necessary" for saddle pts. You may have problems w/o saddle pts or w/o strong duality.

Ex: $\min_x e^{-x}$

Ⓑ only needed if differentiable

Ⓒ If we're nonconvex, not sufficient, ie., there may be non-optimal solutions

Thm 1 restated: optimal solutions satisfy KKT equations

(but not all sol'n of KKT equations are optimal)

Now, add convexity

modify "stationarity" ① to be more general:

① $x^* \in \operatorname{argmin}_x \mathcal{L}(x, \lambda^*, \nu^*)$

subdiff. is w.r.t. x (not λ, ν)

i.e., by convexity, $\boxed{0 \in \partial \mathcal{L}(x^*, \lambda^* \ \nu^*)}$

i.e., if CQ, $0 \in \partial f_0(x^*) + \Sigma' \lambda_i \ \partial f_i(x^*) + \Sigma' \nu_i^* \partial h_i(x^*)$

i.e., if also differentiable,

$0 = \nabla f_0(x^*) + \Sigma' \lambda_i^* \nabla f_i(x^*) + \Sigma' \nu_i^* \nabla h_i(x^*)$

... as before

KKT Thm 2, convexity: sufficient

If (P) is convex (ie., f_i, $i=0,1,\ldots,m$ are convex, $h_i(x) = a_i^T x + b_i$)

then if (x^*, λ^*, ν^*) solve the KKT conditions, then then they are primal/dual optimal and there's no duality gap

proof: Assume (x^*, λ^*, ν^*) satisfy KKT conditions

$p^* \overset{=}{\le} f_0(x^*)$ since x^* is feasible

$= \mathcal{L}(x^*, \lambda^*, \nu^*)$ since feasible + comp. slackness

$= \inf_x \mathcal{L}(x, \lambda^*, \nu^*)$ by stationarity

$= g(\lambda^*, \nu^*)$ by dual feasibility

$\le d^*$ $\implies$ $p^* \le d^*$ (and $d^* \le p^*$) $\implies d^* = p^*$

So, strong duality, and all "$\le$" must "$=$"

So $p^* = f_0(x^*)$ $\implies$ x^* is optimal, same for duals

KKT Thm 3, Necessary AND sufficient (convex case)

If (P) is convex and Slater's conditions hold, then

(x^* is primal optimal) iff (x^* (and some λ^*, ν^*) satisfy the KKT conditions)

proof: Slater's $\Rightarrow$ strong duality and existence of λ^*, ν^*

So either x^* doesn't exist (ex: $\min_x e^{-x}$)

or else it does and so KKT Thms 1 and 2 apply.

Remark In complementary slackness, we said if $f_j(x^*) < 0$ then $\lambda_j^* = 0$ and hence (under strong duality) since

$x^* \in \operatorname{argmin} \underbrace{\mathcal{L}(x, \lambda^*, \nu^*)}$, it's as if the constraint didn't exist

$$= f_0(x) + \sum_{i \neq j} \lambda_i^* f_i(x^*) + \sum_i \nu_i^* h_i(x^*)$$

... but this isn't true if you don't have strong duality.

In particular, it usually isn't true for non-convex problems

Ex: Strong duality / convexity

(P) $\min_{x \in \mathbb{R}} x$ s.t. $x \geq 0$, $\underbrace{x \leq 1}_{\text{not tight}}$

Solution is $x^* = 0$, so $x \leq 1$ inequality had no effect. We could have solved

$\min_{x \in \mathbb{R}} x$ s.t. $x \geq 0$ (dropping $x \leq 1$ constraint)

Ex: no strong duality / nonconvex

(P) $\min_{x \in \mathbb{R}} x$

$x \geq 0$] not tight

$x^2 \geq 1$] non-convex

Solution is $x^* = 1$, so $x \geq 0$ inequality wasn't tight

But... we cannot drop it, since

$\min_{x \in \mathbb{R}} x$, $x^2 \geq 1$ has a different sol'n (i.e., no solution: $x \to -\infty$)