R. Herzog, G. Müller, E. Herberg, M. Hashemi
Heidelberg University

Exercise 10 - Solution

Date issued: 19th June 2023
Date due: 27th June 2023

Homework Problem 10.1 (Solvability and global solutions of equality constrained QPs) 6 Points

Prove Lemma 11.2 of the lecture notes, i. e., the following statements for the quadratic problem

$$\begin{aligned} &\text{Minimize} && \mathcal{L}(\bar{x}, \bar{\lambda}) + \mathcal{L}_x(\bar{x}, \bar{\lambda})\, d + \frac{1}{2} d^\intercal \mathcal{L}_{xx}(\bar{x}, \bar{\lambda})\, d, \quad \text{where } d \in \mathbb{R}^n \\ &\text{subject to} && h(\bar{x}) + h'(\bar{x})\, d = 0. \end{aligned} \tag{11.1}$$

(i) Suppose that the linear system $h(\bar{x}) + h'(\bar{x})\, d = 0$ is solvable, and that d_{part} is some particular solution. Suppose, moreover, that the reduced Hessian $Z^\intercal \mathcal{L}_{xx}(\bar{x}, \bar{\lambda})\, Z$ is positive semidefinite. Then the objective in the reduced QP (11.3) is convex. In this case, the following are equivalent:

(a) The QP (11.1) possesses at least one (global) minimizer.

(b) The QP (11.1) is neither unbounded nor infeasible.

(c) The KKT conditions (11.2) are solvable.

(d) The reduced QP (11.3) possesses at least one (global) minimizer.

(e) The reduced QP (11.3) is not unbounded.

(f) The first-order optimality condition (11.4) is solvable.

The global minimizers of (11.1) are precisely the KKT points, i. e., the d-components of solutions (d, λ) to the KKT system (11.2).

(ii) Suppose that the linear system $h(\overline{x}) + h'(\overline{x})\, d = 0$ is solvable, and that d_{part} is some particular solution. Suppose now that the reduced Hessian $Z^\intercal \mathcal{L}_{xx}(\overline{x}, \overline{\lambda})\, Z$ is not positive semidefinite. Then the QP (11.1) and the reduced QP (11.3) are unbounded.

(iii) Suppose that the linear system $h(\overline{x}) + h'(\overline{x})\, d = 0$ is not solvable. Then the QP (11.1) is infeasible and the reduced QP cannot be formulated for lack of a particular solution d_{part}.

Solution.

(i) Clearly, the reformulation in the lecture notes that used the representation of feasible d as

$$d = d_{\text{part}} + Z\, y$$

shows that the original constrained and the reduced problems are equivalent.

Accordingly, we immediately see that Statement (a) $\Leftrightarrow$ Statement (d) and Statement (b) $\Leftrightarrow$ Statement (e).

Additionally, Lemma 4.1 immediately applies to the reduced problem

$$\text{Minimize} \quad \left[\mathcal{L}_x(\overline{x}, \overline{\lambda}) + d_{\text{part}}^\intercal \mathcal{L}_{xx}(\overline{x}, \overline{\lambda})\right] Z\, y + \frac{1}{2}\, y^\intercal Z^\intercal \mathcal{L}_{xx}(\overline{x}, \overline{\lambda})\, Z\, y, \quad \text{where } y \in \mathbb{R}^{n-r}. \tag{11.3}$$

so we obtain that Statement (d) $\Leftrightarrow$ Statement (e) $\Leftrightarrow$ Statement (f) and that the global minimizers are precisely the solutions to the reduced first order optimality conditions (11.2).

Accordingly, we only need to tie in Statement (c). Since the constraints are affine, we know that ACQ holds at every feasible point (see homework problem 8.4), hence every (local) minimizer is a KKT point making up the d components of the KKT conditions. We would like to show the reverse statement using Theorem 8.19, but that is formulated using the convexity of the cost functional everywhere (which we only have on the feasible set). One could adjust the proof to deal with the weaker assumption but would have to reformulate the lower linear approximation result for convex functions as well, so instead, we proceed to notice that if (d, λ) is a KKT point, then d is feasible and accordingly

$$d = d_{\text{part}} + Z\, y$$

Plugging this into the first line of the block system yields that

$$\mathcal{L}_{xx}(\overline{x}, \overline{\lambda})(d_{\text{part}} + Z\, y) + h'(\overline{x})^\intercal \lambda = -\nabla_x \mathcal{L}(\overline{x}, \overline{\lambda})$$

and, since Z spans the kernel of $h'(\overline{x})$, multiplying this line by $Z^\top$ yields the first order system (11.4), meaning that y is a global minimizer to the reduced QP and therefore d is a global minimizer for the original QP.

(ii) Both the QP and the reduced QP are equivalent. The unboundedness is easiest to see for the reduced problem. Since $Z^\top \mathcal{L}_{xx} Z$ is indefinite, it has at least one negative eigenvalue $\nu < 0$ with corresponding eigenvector $\tilde{y} \neq 0$, i. e., a direction of negative curvature in the feasible set. Accordingly:

$$\begin{aligned}\left[\mathcal{L}_x(\overline{x},\overline{\lambda}) + d_{\text{part}}^\top \mathcal{L}_{xx}(\overline{x},\overline{\lambda})\right] Z t\tilde{y} + \frac{1}{2} t\tilde{y}^\top Z^\top \mathcal{L}_{xx}(\overline{x},\overline{\lambda})\, Z t\tilde{y} =\\ \frac{1}{2} t\nu \|\tilde{y}\|_2^2 + \left[\mathcal{L}_x(\overline{x},\overline{\lambda}) + d_{\text{part}}^\top \mathcal{L}_{xx}(\overline{x},\overline{\lambda})\right] Z t\tilde{y}\\ \xrightarrow{t\to\infty} -\infty\end{aligned}$$

meaning that we can reach arbitrarily small function values by feasible points.

(iii) This is obvious.

(6 Points)

Homework Problem 10.2 (LICQ is equivalent to a unique Lagrange multiplier for certain QPs) 3 Points

Consider the (affine linearly) equality constrained quadratic optimization problem of the form

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2} x^\top A x + b^\top x + c, \quad \text{where } x \in \mathbb{R}^n \\ \text{subject to} \quad & Cx = d \end{aligned} \tag{0.1}$$

for symmetric $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$ and $C \in \mathbb{R}^{n \times n_{\text{eq}}}$, $d \in \mathbb{R}^{n_{\text{eq}}}$ and let x^* be a KKT-point of (0.1).

Show that the set $\Lambda(x^*)$ of Lagrange multipliers corresponding to x^* is a singleton if and only if the LICQ is satisfied at x^*.

Note: This proves the second set of equivalences in Equation (11.5).

Solution.

The KKT system of (0.1) for $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^{n_{\text{eq}}}$ is

$$
\begin{aligned}
Ax + b + C^\mathsf{T}\lambda &= 0 \\
Cx &= d
\end{aligned}
\tag{0.2}
$$

and by assumption, there exists a $\lambda^* \in \mathbb{R}^{n_{\text{eq}}}$ corresponding to x^* such that (x^*, λ^*) solves the system (0.2). Due to the first line of the same system, we know that

$$
\Lambda(x^*) = \lambda^* + \ker C^\mathsf{T}.
$$

Additionally, $\ker C^\mathsf{T} = \{0\}$ if and only if C^T has full column rank, which is exactly the LICQ at x^*. (However, note that for this affine constraint, LICQ is satisfied at all feasible points or at none of them). Otherwise the kernel is a nontrivial subspace of positive dimension, meaning $\Lambda(x^*)$ is a nontrivial affine subspace (and therefore not compact, which we were to expect because LICQ is equivalent to MFCQ for equality constrained problems and MFCQ is equivalent to compact multiplier sets).

(3 Points)

Homework Problem 10.3 (Complementarity is equivalent to variational inequality) 3 Points

Prove Lemma 11.4 of the lecture notes, i. e. the equivalence of the KKT complementarity condition

$$
\mu \ge 0, \quad g(x) \le 0, \quad \mu^\mathsf{T} g(x) = 0 \tag{11.11b}
$$

and the variational inequality

$$
\mu \in K \quad \text{and} \quad g(x)^\mathsf{T}(\nu - \mu) \le 0 \quad \text{for all } \nu \in K \tag{11.12}
$$

with the closed convex cone $K := \mathbb{R}^{n_{\text{ineq}}}_{\ge 0}$ (the non-negative orthant).

Solution.

(3 Points) Suppose that the vectors μ and $g(x)$ satisfy the complementarity system (11.11b). Denote the active and inactive indices at x by $\mathcal{A}(x)$ and $\mathcal{I}(x)$, respectively. We obtain

$$
g(x)^\mathsf{T}(\nu - \mu) = \sum_{i \in \mathcal{A}(x)} \underbrace{g_i(x)}_{=0} (\nu_i - \mu_i) + \sum_{i \in \mathcal{I}(x)} \underbrace{g_i(x)}_{<0} (\underbrace{\nu_i}_{\ge 0} - \underbrace{\mu_i}_{=0}) \le 0
$$

for all $\nu \in K$, i. e., (11.12) holds.

Conversely, suppose that (11.12) is true. For an arbitrary index i, we may have three cases. In case $g_i(x) < 0$, we must have $\mu_i = 0$. Otherwise the choice $\nu = \mu \pm \varepsilon e_i \in K$ (with the i-the standard basis vector e_i) gives a contradiction. In case $g_i(x) = 0$ nothing is to be shown. The case $g_i(x) > 0$ yields a contradiction to (11.12) when we choose $\nu = \mu + e_i$, so it cannot occur. This shows (11.11b).

Homework Problem 10.4 (On the normal cone) 3 Points

Prove Lemma 11.6 of the lecture notes, i. e., the following statements for a set $M \subseteq \mathbb{R}^n$ and $x \in M$:

(i) The normal cone is a closed convex cone.

(ii) $\mathcal{N}_M(x) = (M - \{x\})^\circ$ holds.

Additionally, prove that

(iii) $\mathcal{N}_M(x) \subseteq \mathcal{T}_M(x)^\circ$ but generally $\mathcal{N}_M(x) \neq \mathcal{T}_M(x)^\circ$.

Solution.

(i) This is due to the non-strict defining inequality and the continuity (and linearity) of the scalar product.

(ii) This is a direct consequence of the definition of the polar cone and the Minkowski sum.

(iii) Let $s \in \mathcal{N}_M(x)$ and $d \in \mathcal{T}_M(x)$ with the corresponding sequences $t^{(k)} \searrow 0$ and $x^{(k)} \in M$ such that
$$d^{(k)} := \frac{x^{(k)} - x}{t^{(k)}} \to d.$$
Then
$$s^\intercal d^{(k)} = \frac{\overbrace{s^\intercal (x^{(k)} - x)}^{\le 0}}{\underbrace{t^{(k)}}_{\ge 0}} \le 0$$
and therefore
$$s^\intercal d \le 0$$

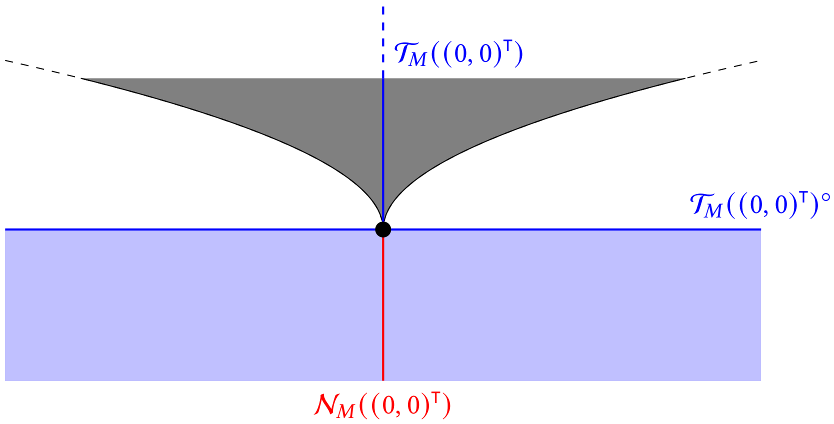

Figure 0.1: Set M (gray), tangent cone with its polar and the normal cone at the origin. The tangent cone and its polar only use local information while the normal cone requires information on the entire set.

due to continuity.

The inverse inclusion generally does not hold because the tangent cone (and hence its polar) only considers *local* information, while the normal cone is built on information concerning the entire set.

Consider, e. g., the set

$$M := \{x \in \mathbb{R}^2 \mid x_2 \geq \sqrt{|x_1|}\}$$

at the origin, where

$$\mathcal{T}_M(x) = 0 \times \mathbb{R}_{\geq} \quad \text{and} \quad \mathcal{T}_M(x)^\circ = \mathbb{R} \times \mathbb{R}_{\leq}$$

but

$$\mathcal{N}_M(x) = 0 \times \mathbb{R}_{\leq} \subsetneq \mathcal{T}_M(x)^\circ.$$

(3 Points)

Please submit your solutions as a single pdf and an archive of programs via moodle.