

Regularized Equality Constrained Quadratic Optimization

Anton Xue and Nikolai Matni

1 Introduction

We look at regularized approximations of equality-constrained quadratic programming. In particular, what is $\|x^\star - x_\lambda^\star\|$, where x^\star is the optimal solution to the original problem, while x_λ^\star is the solution to the regularized problem.

2 Background

Stephen Boyd and Lieven Vandenberghe [1].

2.1 Equality Constrained Quadratic Programming

A quadratic program with equality constraints [1] is

$$\text{minimize} \quad \frac{1}{2}x^\top Qx \tag{1}$$

$$\text{subject to} \quad Ax = b \tag{2}$$

with variable in $x \in \mathbb{R}^n$, where $Q \succeq 0$ and the constraint $A \in \mathbb{R}^{m \times n}$ is, for our purposes, fat and full rank.

The ℓ^2 regularized version for $\lambda > 0$ looks like

$$\text{minimize} \quad \frac{1}{2}x^\top Qx + \lambda x^\top x \tag{3}$$

$$\text{subject to} \quad Ax = b \tag{4}$$

2.2 Karush-Kuhn-Tucker Conditions

The Lagrangian for (1) is

$$L(x, p) = \frac{1}{2}x^\top Qx + p^\top (Ax - b)$$

for which the optimality conditions [1] are

$$Qx^* + A^\top p^* = b, \quad Ax^* = b$$

or more compcatctly expressed:

$$\begin{bmatrix} Q & A^\top \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ p^* \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix} \quad (5)$$

On the other hand the Lagrangian of the regularized problem (3) is

$$L_\lambda(x, p) = \frac{1}{2}x^\top (Q + \lambda I)x + p^\top (Ax - b)$$

which has the KKT conditions

$$\begin{bmatrix} Q + \lambda I & A^\top \\ A & 0 \end{bmatrix} \begin{bmatrix} x_\lambda^* \\ p_\lambda^* \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix} \quad (6)$$

2.3 Matrix Inversion

Consider a symmetric matrix

$$Q = \begin{bmatrix} A & B \\ B^\top & C \end{bmatrix}$$

when $A \succ 0$, define the Schur complement $S = C - B^\top A^{-1}B$. Then

$$\begin{bmatrix} A & B \\ B^\top & C \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}BS^{-1}B^\top A^{-1} & -A^{-1}BS^{-1} \\ -S^{-1}B^\top A^{-1} & S^{-1} \end{bmatrix}$$

3 Bounding Norms

Noting that by linearity

$$\begin{bmatrix} Q & A^\top \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ p^* \end{bmatrix} + \begin{bmatrix} \lambda I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x^* \\ p^* \end{bmatrix} = \begin{bmatrix} Q + \lambda I & A^\top \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ p^* \end{bmatrix} = \begin{bmatrix} \lambda x^* \\ b \end{bmatrix}$$

and so through subtracting equations,

$$\begin{bmatrix} Q + \lambda I & A^\top \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* - x_\lambda^* \\ p^* - p_\lambda^* \end{bmatrix} = \begin{bmatrix} \lambda x^* \\ 0 \end{bmatrix}$$

Given our assumptions on $Q + \lambda I \succ 0$ and A is full rank. For ease of notation, define $\Lambda = Q + \lambda I$. The Schur complement is then $S = -\Lambda^{-1}A^\top$, and

$$\begin{bmatrix} x^* - x_\lambda^* \\ p^* - p_\lambda^* \end{bmatrix} = \begin{bmatrix} \Lambda^{-1} + \Lambda^{-1}A^\top S^{-1}A\Lambda^{-1} & -\Lambda^{-1}A^\top S^{-1} \\ -S^{-1}A\Lambda^{-1} & S^{-1} \end{bmatrix} \begin{bmatrix} \lambda x^* \\ 0 \end{bmatrix}$$

In other words:

$$x^* - x_\lambda^* = (\Lambda^{-1} - \Lambda^{-1}A^\top(A\Lambda^{-1}A^\top)^{-1}A\Lambda^{-1})\lambda x^*$$

To simplify notation slightly, use $\Gamma = \Lambda^{-1}$ because it looks like an upside-down L . Then

$$x^* - x_\lambda^* = (\Gamma - \Gamma A^\top(A\Gamma A^\top)^{-1}A\Gamma)\lambda x^*$$

Theorem 1. *Without loss of generality, assume that*

$$Q = \text{diag}(q_1, q_2, \dots, q_n), \quad q_1 \geq \dots \geq q_n \geq 0.$$

Then for $\lambda \leq q_1$,

$$\|x^* - x_\lambda^*\| \leq \frac{q_1 - q_n}{(q_1 + \lambda)^2} \lambda \|x^*\|$$

Proof. Our angle of attack is to seek a bound $\|\Gamma - \Gamma A^\top(A\Gamma A^\top)^{-1}A\Gamma\|$. For this, note that both $\Gamma \succ 0$ and $\Gamma A^\top(A\Gamma A^\top)^{-1}A\Gamma \succ 0$ and so one technique is to find a matrix G such that

$$G \prec \Gamma A^\top(A\Gamma A^\top)^{-1}A\Gamma \implies \|\Gamma - \Gamma A^\top(A\Gamma A^\top)^{-1}A\Gamma\| \leq \|\Gamma - G\|$$

We can iteratively construct G . Let a singular value decomposition of A be

$$A = U\Sigma V^\top = U \begin{bmatrix} \Sigma_1 & 0 \end{bmatrix} \begin{bmatrix} V_1^\top \\ V_2^\top \end{bmatrix} = U\Sigma_1 V_1^\top, \quad \Sigma_1 = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_m)$$

Then first upper-bounding the inside of the inverse in order to lower-bound the overall inverse,

$$A\Gamma A^\top = U\Sigma_1 V_1^\top \begin{bmatrix} \frac{1}{q_1 + \lambda} & & \\ & \ddots & \\ & & \frac{1}{q_n + \lambda} \end{bmatrix} V_1 \Sigma_1 U^\top \preceq \frac{1}{q_n + \lambda} U\Sigma_1^2 U^\top$$

Then for the inverse we would have

$$\begin{aligned} \Gamma A^\top(A\Gamma A^\top)^{-1}A\Gamma &\preceq \Gamma A^\top[(q_n + \lambda)U\Sigma_1^{-2}U^\top]A\Gamma \\ &= (q_n + \lambda)\Gamma V_1 \Sigma_1 U^\top [U\Sigma_1^{-2}U^\top] U\Sigma_1 V_1^\top \Gamma \\ &\preceq (q_n + \lambda)\Gamma \Gamma \end{aligned}$$

Putting these together, when $\lambda \leq q_1$

$$\begin{aligned} \Gamma - \Gamma A^\top(A\Gamma A^\top)^{-1}A\Gamma &\preceq \Gamma - (q_n + \lambda)\Gamma \Gamma \\ &\preceq \begin{bmatrix} \frac{1}{q_1 + \lambda} & & \\ & \ddots & \\ & & \frac{1}{q_n + \lambda} \end{bmatrix} - (q_n + \lambda) \begin{bmatrix} \frac{1}{(q_1 + \lambda)^2} & & \\ & \ddots & \\ & & \frac{1}{(q_n + \lambda)^2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{q_1 + \lambda}{(q_1 + \lambda)^2} - \frac{q_n + \lambda}{(q_1 + \lambda)^2} & & \\ & \ddots & \\ & & \frac{q_n + \lambda}{(q_n + \lambda)^2} - \frac{q_n + \lambda}{(q_n + \lambda)^2} \end{bmatrix} \preceq \frac{q_1 - q_n}{(q_1 + \lambda)^2} I \end{aligned}$$

Consequently, for the original bound

$$\begin{aligned}
\|x^* - x_\lambda^*\| &= \|(\Gamma - \Gamma A^\top (A\Gamma A^\top)^{-1} A\Gamma) \lambda x^*\| \\
&\leq \|\Gamma - \Gamma A^\top (A\Gamma A^\top)^{-1} A\Gamma\| \cdot \|\lambda x^*\| \\
&\leq \frac{q_1 - q_n}{(q_1 + \lambda)^2} \lambda \|x^*\|
\end{aligned}$$

□

References

- [1] Stephen P Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.