Homework 2 Supplementary Note

Spring 2019
Stanley Chan

Here are some hints about the last problem. First let us define

$$\begin{aligned} \boldsymbol{x}_\lambda &= \underset{\boldsymbol{x}}{\operatorname{argmin}} \quad \|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{y}\|^2 + \lambda\|\boldsymbol{x}\|^2 \\ &= (\boldsymbol{A}^T\boldsymbol{A}+\lambda\boldsymbol{I})^{-1}\boldsymbol{A}^T\boldsymbol{y}, \end{aligned}$$

where the inverse exists if we assume $\boldsymbol{A}$ is full rank.

Now, let us look at the constrained problem

$$\boldsymbol{x}_\alpha = \underset{\boldsymbol{x}}{\operatorname{argmin}} \quad \|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{y}\|^2 \quad \text{subject to} \quad \|\boldsymbol{x}\|^2 \leq \alpha.$$

We ask: When will $\boldsymbol{x}_\alpha = \boldsymbol{x}_\lambda$? And is this $\boldsymbol{x}_\alpha$ the optimizer of the constrained problem?

The first thing you need to do is pick a candidate $\boldsymbol{x}_\alpha$. We can identify this candidate by looking at the Lagrangian function:

$$L(\boldsymbol{x}_\alpha, \nu) = \|\boldsymbol{A}\boldsymbol{x}_\alpha - \boldsymbol{y}\|^2 - \nu(\alpha - \|\boldsymbol{x}_\alpha\|^2).$$

- The stationarity condition suggests that

  $$\boldsymbol{0} = \nabla_{\boldsymbol{x}_\alpha} L(\boldsymbol{x}_\alpha, \nu) = 2\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{x}_\alpha - \boldsymbol{y}) + 2\nu\boldsymbol{x}_\alpha,$$

  which implies

  $$\boldsymbol{x}_\alpha = (\boldsymbol{A}^T\boldsymbol{A}+\nu\boldsymbol{I})^{-1}\boldsymbol{A}^T\boldsymbol{y}.$$

  When will $\boldsymbol{x}_\alpha = \boldsymbol{x}_\lambda$? We can make this happen when we let $\nu = \lambda$. So we have found our candidate: Make $\nu = \lambda$, then $\boldsymbol{x}_\alpha = \boldsymbol{x}_\lambda$. This $\boldsymbol{x}_\alpha$ satisfies the stationarity condition.

- Now, let us look at the primal feasibility. We need

  $$\|\boldsymbol{x}_\alpha\|^2 \leq \alpha.$$

  To make this happen, we choose $\alpha = \|\boldsymbol{x}_\lambda\|^2$. Then the primal feasibility is guaranteed to work because we just chose $\boldsymbol{x}_\alpha = \boldsymbol{x}_\lambda$.

- What about the dual feasibility? We need $\nu \geq 0$. But this is automatically given because we chose $\nu = \lambda$.

- Finally the complementary slackness. We need $\nu(\alpha - \|\boldsymbol{x}_\alpha\|^2) = 0$. By primal feasibility, we have $\alpha = \|\boldsymbol{x}_\lambda\|^2$, and so the complementary slackness is okay.

To conclude: If we are given $\boldsymbol{x}_\lambda$, we can make $\boldsymbol{x}_\alpha = \boldsymbol{x}_\lambda$ by choosing $\alpha = \|\boldsymbol{x}_\lambda\|^2$ and $\nu = \lambda$.

Remark: This result shows that the solution is always found at the boundary of the circle formed by $\|\boldsymbol{x}\|^2 = \alpha$ and the ellipse formed by $\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{y}\|^2$. You will never go into the interior of the constraint set. If you ever go into the interior, it means that the α you choose is not matching with the λ you need to show the equivalence $\boldsymbol{x}_\alpha = \boldsymbol{x}_\lambda$.