Solutions to Part A of Problem Sheet 7

Solution (7.1) First of all, note that the function is only defined for $\boldsymbol{x}$ such that $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$. This is the *domain* of the function.

We introduce new variables $\boldsymbol{y}$ and derive the dual to the problem

$$\begin{aligned} \text{minimize} \quad & -\sum_{i=1}^{m} \log(y_i) \\ \text{subject to} \quad & \boldsymbol{y} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}. \end{aligned}$$

Note that by restricting to the domain of the problem, we don't have to explicitly ask for $\boldsymbol{y}$ to be non-negative: the objective function wouldn't make sense for negative values.

The Lagrangian to this problem is

$$\begin{aligned} \mathcal{L}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\mu}) &= -\sum_{i=1}^{m} \log(y_i) + \boldsymbol{\mu}^\top (\boldsymbol{y} - \boldsymbol{b} + \boldsymbol{A}\boldsymbol{x}) \\ &= \sum_{i=1}^{m} -\log(y_i) + \mu_i (y_i - b_i + \boldsymbol{a}_i^\top \boldsymbol{x}). \end{aligned}$$

The dual function is

$$g(\boldsymbol{\mu}) = \inf_{\boldsymbol{x}, \boldsymbol{y}} \mathcal{L}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\mu}),$$

where the infimum is taken over the domain of $\mathcal{L}$ (in particular, this requires $\boldsymbol{y} \geq \boldsymbol{0}$). The infimum is $-\infty$ if $\boldsymbol{\mu}^\top \boldsymbol{A} \neq \boldsymbol{0}$. If $\boldsymbol{\mu}$ has negative terms, then the infimum is also $-\infty$ (we could then choose an arbitrary large value for the corresponding y variable).

If $\boldsymbol{\mu} > 0$, then we can determine the minimum by computing the gradient. For the partial derivative in y_i we get

$$\frac{\partial \mathcal{L}}{\partial y_i} = -\frac{1}{y_i} + \mu_i = 0,$$

so at the minimum we have $y_i = \frac{1}{\mu_i}$. For the gradient in the $\boldsymbol{x}$ variables we get $\nabla_{\boldsymbol{x}} \mathcal{L} = \boldsymbol{A}^\top \boldsymbol{\mu} = \boldsymbol{0}$. It follows that the dual function is

$$g(\boldsymbol{\mu}) = \begin{cases} \sum_{i=1}^{m} \log(\mu_i) + m - \boldsymbol{b}^\top \boldsymbol{\mu} & \text{if } \boldsymbol{A}^\top \boldsymbol{\mu} = \boldsymbol{0},\ \boldsymbol{\mu} > 0, \\ -\infty & \text{else,} \end{cases}$$

where we used that $\log(y_i) = \log(1/\mu_i) = -\log(\mu_i)$.

Solution (7.2) The problem is not convex since the equality constraint is not linear, and the inequality constraint is not convex. We can formulate an equivalent convex optimization problem as

$$\text{minimize} \quad x_1^2 + x_2^2 \qquad \text{subject to} \quad x_1 \leq 0,\ x_1 + x_2 = 0.$$

Solution (7.3) Write

$$\boldsymbol{P}(\lambda) = \boldsymbol{P} + \sum_{i=1}^{m} \lambda_i \boldsymbol{P}_i,\ \boldsymbol{q}(\lambda) = \boldsymbol{q} + \sum_{i=1}^{m} \lambda_i \boldsymbol{q}_i,\ \boldsymbol{r}(\lambda) = \boldsymbol{r} + \sum_{i=1}^{m} \lambda_i \boldsymbol{r}_i.$$

With this notation, we can express the Lagrangian as

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{P}(\lambda)\boldsymbol{x} + \boldsymbol{q}(\lambda)^\top \boldsymbol{x} + \boldsymbol{r}(\lambda).$$

We can now approach this minimization problem just as we would approach any such problem with a positive semidefinite matrix: compute the gradient in $\boldsymbol{x}$, $\boldsymbol{P}(\boldsymbol{\lambda}) + \boldsymbol{q}(\boldsymbol{\lambda})$, and set this to zero. Plugging in the result, $\boldsymbol{x} = -\boldsymbol{P}(\boldsymbol{\lambda})^{-1}\boldsymbol{q}(\boldsymbol{\lambda})$, into the equation for the Lagrangian, we get for $\boldsymbol{\lambda} \geq \boldsymbol{0}$

$$g(\boldsymbol{\lambda}) = \inf_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = -\frac{1}{2}\boldsymbol{q}(\boldsymbol{\lambda})^\top \boldsymbol{P}(\boldsymbol{\lambda})^{-1}\boldsymbol{q}(\boldsymbol{\lambda}) + \boldsymbol{r}(\boldsymbol{\lambda}).$$

The Lagrange dual is then given by

$$\begin{array}{ll} \text{maximize} & -\dfrac{1}{2}\boldsymbol{q}(\boldsymbol{\lambda})^\top \boldsymbol{P}(\boldsymbol{\lambda})^{-1}\boldsymbol{q}(\boldsymbol{\lambda}) + \boldsymbol{r}(\boldsymbol{\lambda}) \\ \text{subject to} & \boldsymbol{\lambda} \geq \boldsymbol{0}. \end{array}$$

This function looks simpler at first sight, since it only involves non-negativity constraints, but it requires the inverse of a linear combination of the matrices $\boldsymbol{P}_i$, which makes things less straight-forward.