EE 150 - Applications of Convex Optimization in Signal Processing and Communications
Dr. Andre Tkacenko, JPL
Third Term 2011-2012

Homework Set #6 - Solutions

1. (a) The feasible set is the interval $[2, 4]$. The (unique) optimal point or solution is $x^\star = 2$, and the optimal value is $p^\star = 5$.

(b) Here, we have $f_0(x) = x^2 + 1$ and $f_1(x) = (x-2)(x-4) = x^2 - 6x + 8$. Thus, the Lagrangian $L(x, \lambda)$ is given by

$$L(x, \lambda) = f_0(x) + \lambda f_1(x) = (\lambda + 1) x^2 - 6\lambda x + (8\lambda + 1) \; .$$

A plot of the objective function, the constraint function, the feasible set, the optimal point and value, and the Lagrangian $L(x, \lambda)$ for several positive values of λ is shown in Figure 1. From the plot, it is clear that minimum value of $L(x, \lambda)$ over x (i.e., $g(\lambda)$) is always less than $p^\star$. It increases as λ varies from 0 to 2, reaches its maximum at $\lambda = 2$, and then decreases again as λ increases above 2. We have equality, namely $p^\star = g(\lambda)$ for $\lambda = 2$.

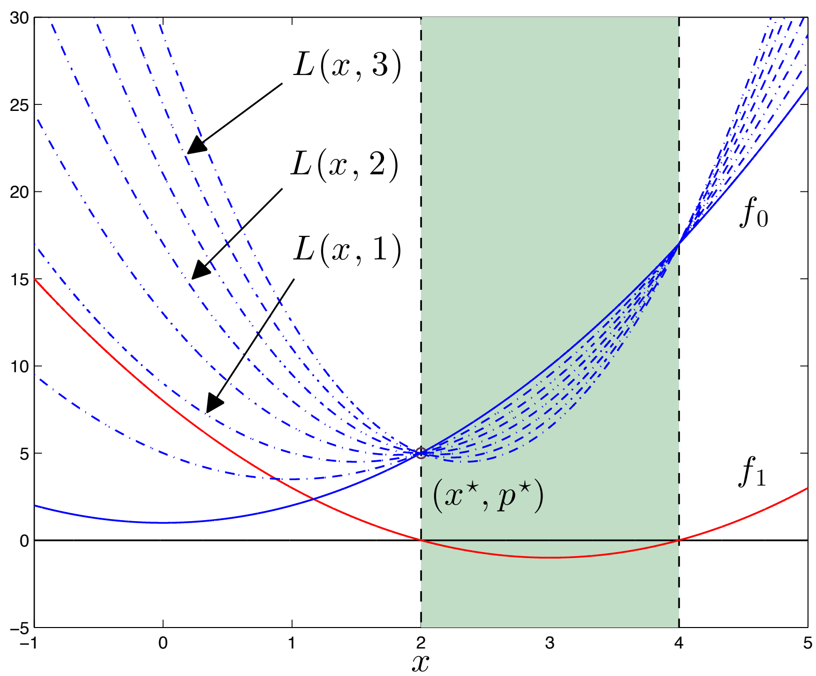


Figure 1: Plot of the objective function $f_0(x)$, the constraint function $f_1(x)$, the feasible set $[2, 4]$, the optimal point and value $(x^\star, p^\star) = (2, 5)$, along with Lagrangian $L(x, \lambda)$ for several positive values of λ.

For $\lambda > 1$, the Lagrangian is a convex parabola and reaches its minimum at $\widetilde{x} = 3\lambda / (1 + \lambda)$. On the other hand, for $\lambda \leq -1$, the Lagrangian is a concave parabola and as such is unbounded below. Thus, the dual function $g(\lambda)$ is given by

$$g(\lambda) = \begin{cases} -\dfrac{9\lambda^2}{\lambda + 1} + 8\lambda + 1 \, , & \lambda > 1 \\ -\infty \, , & \lambda \leq -1 \end{cases}$$

A plot of the dual function $g(\lambda)$ is shown in Figure 2.

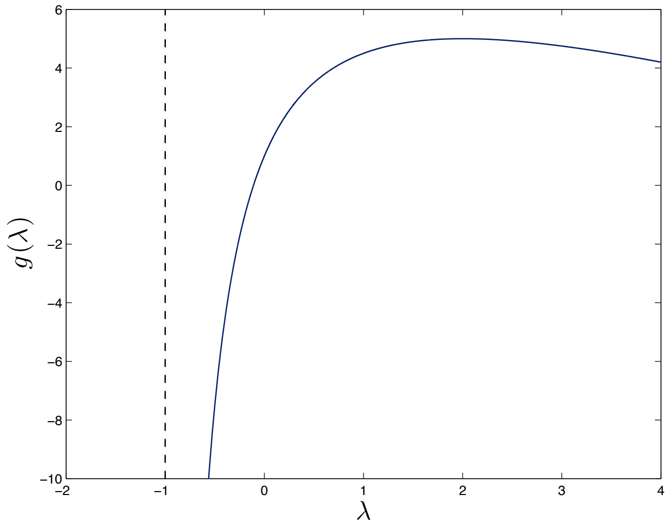


Figure 2: Plot of the dual function $g(\lambda)$ versus λ.

We can verify that the dual function is concave, that its value is equal to $p^\star = 5$ for $\lambda = 2$, and less than $p^\star$ for other values of λ. Specifically, by taking the derivative of $g(\lambda)$ with respect to λ and setting the result to zero leads to the condition

$$(\lambda - 2)\,(\lambda + 4) = 0\,,$$

which for $\lambda > -1$ can only be satisfied for $\lambda = 2$. For $\lambda = 2$, we have $g(2) = 5 = p^\star$ indeed.

(c) The Lagrange dual problem is

$$\begin{array}{ll} \text{maximize} & -\dfrac{9\lambda^2}{\lambda+1} + 8\lambda + 1 \\ \text{subject to} & \lambda \geq 0 \end{array}\,.$$

A second derivative test on the objective $g(\lambda)$ yields

$$g''(\lambda) = -\frac{18}{(\lambda+1)^3} < 0\,,\ \forall\, \lambda > -1\,.$$

Thus, the objective is a *concave* function of λ over the domain of the problem and so the dual problem is indeed a concave maximization problem. Setting the derivative of the objective with respect to λ equal to zero yields the condition

$$(\lambda - 2)\,(\lambda + 4) = 0\,,$$

which for $\lambda \geq 0$ can only be satisfied for $\lambda = 2$. We have $g(2) = d^\star = 5$ here. Hence, the dual optimal solution is $\lambda^\star = 2$ and the dual optimal value is $d^\star = 5$. As $p^\star = 5$, we can verify that strong duality indeed holds for this example (as it must, since Slater's constraint qualification is satisfied).

(d) The perturbed problem is infeasible for $u < -1$, since the infimum of $f_1(x)$ about all $x \in \mathbb{R}$ is equal to -1. For $u \geq -1$, the feasible set is the interval

$$\left[3 - \sqrt{u+1}, 3 + \sqrt{u+1}\right] ,$$

given by the two roots of $x^2 - 6x + 8 = u$. For $-1 \leq u \leq 8$, the optimum solution is $x^\star(u) = 3 - \sqrt{1+u}$. On the other hand, for $u > 8$, the optimum is the unconstrained minimum of f_0, i.e., $x^\star(u) = 0$. In summary, we have the following.

$$x^\star(u) = \begin{cases} \text{undefined}\,, & u < -1 \\ 3 - \sqrt{u+1}\,, & -1 \leq u \leq 8 \\ 0\,, & u > 8 \end{cases} \Longrightarrow p^\star(u) = \begin{cases} \infty\,, & u < -1 \\ 11 + u - 6\sqrt{u+1}\,, & -1 \leq u \leq 8 \\ 1\,, & u > 8 \end{cases} .$$

A plot of the optimal value function $p^\star(u)$ and its epigraph is shown in Figure 3. In addition, a supporting hyperplane (or more appropriately, a line in this case) at $u = 0$ is shown.

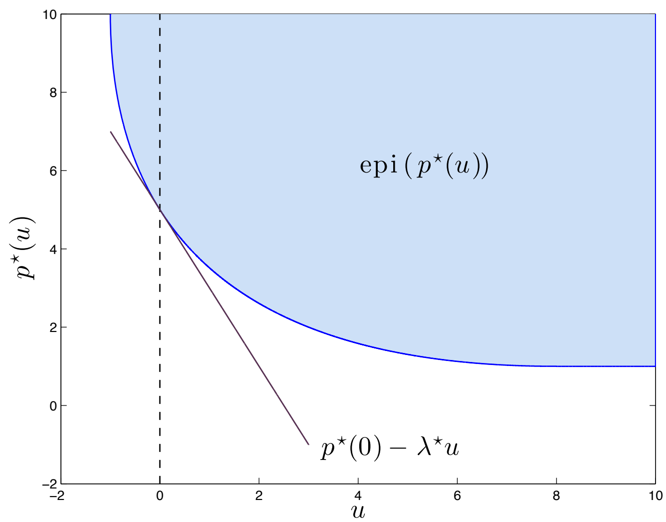


Figure 3: Plot of the optimal value $p^\star(u)$ as a function of the perturbation u. Included with this is a plot of the epigraph of the optimal value, along with a supporting hyperplane at $u = 0$ (corresponding to the nominal optimal value).

From our expression for $p^\star(u)$ above, it can be seen that it is a differentiable function of u at $u = 0$. Note that we have

$$\frac{dp^\star(u)}{du} = 1 - \frac{3}{\sqrt{u+1}} , \quad \text{for } -1 \leq u \leq 8 .$$

Thus, we have

$$\frac{dp^\star(0)}{du} = -2 = -\lambda^\star .$$

2. (a) Consider the objective function of the LP, namely $\mathbf{x}^T\mathbf{y}$. Note that we have the following.

$$\mathbf{x}^T\mathbf{y} = \sum_{k=1}^{n} x_k y_k = \sum_{k=1}^{n} x_{[k]} y_{i_k} ,$$

where i_k for $k = 1, \ldots, n$ is a permutation of the index set $\{1, \ldots, n\}$ corresponding to the elements of $\mathbf{x}$ arranged in descending order. In other words, we have $x_{i_k} = x_{[k]}$ for all k. Under the constraints that $0 \le y_k \le 1$ and $\sum_{k=1}^{n} y_k = r$, it is clear that we have

$$\mathbf{x}^T\mathbf{y} = \sum_{k=1}^{r} x_{[k]} y_{i_k} + \sum_{k=r+1}^{n} x_{[k]} y_{i_k} \le \sum_{k=1}^{r} x_{[k]} = f(\mathbf{x}) \ .$$

In other words, the upper bound on $\mathbf{x}^T\mathbf{y}$ subject to the constraints on $\mathbf{y}$ is obtained by extracting the r largest components of $\mathbf{x}$ and assigning the maximum possible weight of one to each of these components. But this bound can be achieved by setting $y_{i_1} = \cdots = y_{i_r} = 1$ and $y_{i_{r+1}} = \cdots = y_{i_n} = 0$. Hence, $f(\mathbf{x})$ is equal to the optimal value of the LP

$$\begin{array}{ll} \text{maximize} & \mathbf{x}^T\mathbf{y} \\ \text{subject to} & \mathbf{0} \preceq \mathbf{y} \preceq \mathbf{1} \\ & \mathbf{1}^T\mathbf{y} = r \end{array} ,$$

with $\mathbf{y} \in \mathbb{R}^n$ as the variable. This can be more compactly written as

$$f(\mathbf{x}) = \sup\left\{\mathbf{x}^T\mathbf{y} : \mathbf{0} \preceq \mathbf{y} \preceq \mathbf{1}\,,\ \mathbf{1}^T\mathbf{y} = r\right\} \ .$$

(b) First change the objective from maximization to minimization as follows:

$$\begin{array}{ll} \text{minimize} & -\mathbf{x}^T\mathbf{y} \\ \text{subject to} & \mathbf{0} \preceq \mathbf{y} \preceq \mathbf{1} \\ & \mathbf{1}^T\mathbf{y} = r \end{array} .$$

Let us introduce a Lagrange multiplier $\boldsymbol{\lambda}$ for the lower bound on $\mathbf{y}$ (i.e., $\mathbf{y} \succeq \mathbf{0}$), $\mathbf{u}$ for the upper bound on $\mathbf{y}$ (i.e., $\mathbf{y} \preceq \mathbf{1}$), and t for the equality constraint (namely $\mathbf{1}^T\mathbf{y} = r$). This yields the Lagrangian

$$\begin{aligned} L(\mathbf{y}, \boldsymbol{\lambda}, \mathbf{u}, t) &= -\mathbf{x}^T\mathbf{y} - \boldsymbol{\lambda}^T\mathbf{y} + \mathbf{u}^T(\mathbf{y} - \mathbf{1}) + t\left(\mathbf{1}^T\mathbf{y} - r\right) , \\ &= -rt - \mathbf{1}^T\mathbf{u} + (-\mathbf{x} - \boldsymbol{\lambda} + \mathbf{u} + t\mathbf{1})^T\mathbf{y} \,. \end{aligned}$$

Minimizing over $\mathbf{y}$ yields the dual function

$$g(\boldsymbol{\lambda}, \mathbf{u}, t) = \begin{cases} -rt - \mathbf{1}^T\mathbf{u}\,, & -\mathbf{x} - \boldsymbol{\lambda} + \mathbf{u} + t\mathbf{1} = \mathbf{0} \\ -\infty\,, & \text{otherwise} \end{cases} .$$

The dual problem is to maximize g subject to $\boldsymbol{\lambda} \succeq \mathbf{0}$ and $\mathbf{u} \succeq \mathbf{0}$. This yields the problem

$$\begin{array}{ll} \text{maximize} & -rt - \mathbf{1}^T\mathbf{u} \\ \text{subject to} & \boldsymbol{\lambda} = t\mathbf{1} + \mathbf{u} - \mathbf{x} \\ & \boldsymbol{\lambda} \succeq \mathbf{0}\,,\ \mathbf{u} \succeq \mathbf{0} \end{array} .$$

By changing the objective to minimization (by minimizing the negative of the objective above) and eliminating the slack variable $\boldsymbol{\lambda}$ in the equality constraint, we obtain the problem

$$\begin{array}{ll} \text{minimize} & rt + \mathbf{1}^T\mathbf{u} \\ \text{subject to} & t\mathbf{1} + \mathbf{u} \succeq \mathbf{x} \\ & \mathbf{u} \succeq \mathbf{0} \end{array} ,$$

as desired.

(c) The constraint that no more than half of the total power is in any m_0 lamps is equivalent to the condition

$$\sum_{k=1}^{m_0} p_{[k]} \leq \frac{1}{2} \sum_{k=1}^{m} p_k = \frac{1}{2}\mathbf{1}^T\mathbf{p}\,.$$

In other words, if $n \to m$ and $r = m_0$, then the constraint is equivalent to

$$f(\mathbf{p}) \leq \alpha\,,$$

where $\alpha \triangleq \frac{1}{2}\mathbf{1}^T\mathbf{p}$. But from the results of part (b), we know that this condition is true if and only if there exist $s \in \mathbb{R}$ and $\mathbf{q} \in \mathbb{R}^m$ such that

$$m_0 s + \mathbf{1}^T\mathbf{q} \leq \alpha\,,\ \ s\mathbf{1} + \mathbf{q} \succeq \mathbf{p}\,,\ \ \mathbf{q} \succeq \mathbf{0}\,.$$

Thus, with this constraint in effect, the patch illumination problem becomes

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \mathbf{a}_k^T\mathbf{p} \leq I_{\text{des}} t\,,\ \ k = 1, \ldots, n \\ & \left\| \begin{bmatrix} 2\sqrt{I_{\text{des}}} \\ t - \mathbf{a}_k^T\mathbf{p} \end{bmatrix} \right\|_2 \leq t + \mathbf{a}_k^T\mathbf{p}\,,\ \ k = 1, \ldots, n \\ & \mathbf{0} \preceq \mathbf{p} \preceq p_{\max}\mathbf{1} \\ & \alpha = \dfrac{1}{2}\mathbf{1}^T\mathbf{p} \\ & m_0 s + \mathbf{1}^T\mathbf{q} \leq \alpha\,,\ \ s\mathbf{1} + \mathbf{q} \succeq \mathbf{p}\,,\ \ \mathbf{q} \succeq \mathbf{0} \end{array}\,,$$

with variables $\mathbf{t} \in \mathbb{R}$, $\mathbf{p} \in \mathbb{R}^m$, $\alpha \in \mathbb{R}$, $s \in \mathbb{R}$, and $\mathbf{q} \in \mathbb{R}^m$. Eliminating the trivial equality constraint $\alpha = \frac{1}{2}\mathbf{1}^T\mathbf{p}$, we obtain

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \mathbf{a}_k^T\mathbf{p} \leq I_{\text{des}} t\,,\ \ k = 1, \ldots, n \\ & \left\| \begin{bmatrix} 2\sqrt{I_{\text{des}}} \\ t - \mathbf{a}_k^T\mathbf{p} \end{bmatrix} \right\|_2 \leq t + \mathbf{a}_k^T\mathbf{p}\,,\ \ k = 1, \ldots, n \\ & \mathbf{0} \preceq \mathbf{p} \preceq p_{\max}\mathbf{1} \\ & m_0 s + \mathbf{1}^T\mathbf{q} \leq \dfrac{1}{2}\mathbf{1}^T\mathbf{p}\,,\ \ s\mathbf{1} + \mathbf{q} \succeq \mathbf{p}\,,\ \ \mathbf{q} \succeq \mathbf{0} \end{array}\,,$$

with variables $\mathbf{t} \in \mathbb{R}$, $\mathbf{p} \in \mathbb{R}^m$, $s \in \mathbb{R}$, and $\mathbf{q} \in \mathbb{R}^m$, as desired.

3. Introducing $\boldsymbol{\lambda}$ for the inequality constraint $\mathbf{x} \succeq \mathbf{0}$, ν for the scalar equality constraint $\mathbf{1}^T\mathbf{x} = 1$, and $\mathbf{z}$ for the vector equality constraint $\mathbf{P}\mathbf{x} = \mathbf{y}$, the Lagrangian becomes

$$\begin{aligned} L(\mathbf{x}, \mathbf{y}, \boldsymbol{\lambda}, \nu, \mathbf{z}) &= \mathbf{c}^T\mathbf{x} + \sum_{i=1}^{m} y_i \log y_i - \boldsymbol{\lambda}^T\mathbf{x} + \nu\left(\mathbf{1}^T\mathbf{x} - 1\right) + \mathbf{z}^T\left(\mathbf{P}\mathbf{x} - \mathbf{y}\right)\,, \\ &= \left(\mathbf{c} - \boldsymbol{\lambda} + \nu\mathbf{1} + \mathbf{P}^T\mathbf{z}\right)^T\mathbf{x} + \sum_{i=1}^{m} y_i \log y_i - \mathbf{z}^T\mathbf{y} - \nu\,. \end{aligned}$$

The minimum over $\mathbf{x}$ is bounded below if and only if

$$\mathbf{c} - \boldsymbol{\lambda} + \nu\mathbf{1} + \mathbf{P}^T\mathbf{z} = \mathbf{0}\,.$$

To minimize over $\mathbf{y}$, we set the derivative with respect to y_i equal to zero, which gives

$$1 + \log y_i^\star - z_i = 0 \Longleftrightarrow y_i^\star = e^{z_i - 1}\,.$$

Thus, the dual function is

$$\begin{aligned} g(\boldsymbol{\lambda}, \nu, \mathbf{z}) &= \begin{cases} \sum_{i=1}^m e^{z_i-1}(z_i - 1) - \sum_{i=1}^m z_i e^{z_i - 1} - \nu\,, & \mathbf{c} - \boldsymbol{\lambda} + \nu\mathbf{1} + \mathbf{P}^T\mathbf{z} = \mathbf{0} \\ -\infty & \text{otherwise} \end{cases}, \\ &= \begin{cases} -\sum_{i=1}^m e^{z_i-1} - \nu\,, & \boldsymbol{\lambda} = \mathbf{P}^T\mathbf{z} + \nu\mathbf{1} + \mathbf{c} \\ -\infty & \text{otherwise} \end{cases}. \end{aligned}$$

Hence, the dual problem is

$$\begin{array}{ll} \text{maximize} & -\sum_{i=1}^m e^{z_i - 1} - \nu \\ \text{subject to} & \mathbf{P}^T\mathbf{z} + \nu\mathbf{1} + \mathbf{c} \succeq \mathbf{0} \end{array}.$$

This can be simplified by introducing a new variable $\mathbf{w} = \mathbf{z} + \nu\mathbf{1}$ and exploiting the fact that $\mathbf{1} = \mathbf{P}^T\mathbf{1}$. Specifically, the dual problem becomes

$$\begin{array}{ll} \text{maximize} & -\sum_{i=1}^m e^{w_i - \nu - 1} - \nu \\ \text{subject to} & \mathbf{P}^T\mathbf{w} \succeq -\mathbf{c} \end{array}.$$

Finally, we can maximize the objective function with respect to ν by setting the derivative equal to zero. This yields

$$\nu^\star = \log\left(\sum_{i=1}^m e^{w_i - 1}\right).$$

From this, the objective function becomes

$$e^{-\nu^\star}\left(-\sum_{i=1}^m e^{w_i-1}\right) - \nu^\star = -1 - \log\left(\sum_{i=1}^m e^{w_i - 1}\right) = -\log\left(\sum_{i=1}^m e^{w_i}\right).$$

As such, the dual problem can be simplified to

$$\begin{array}{ll} \text{maximize} & -\log\left(\sum_{i=1}^m e^{w_i}\right) \\ \text{subject to} & \mathbf{P}^T\mathbf{w} \succeq -\mathbf{c} \end{array},$$

which can be written as

$$\begin{array}{ll} \text{minimize} & \log\left(\sum_{i=1}^m e^{w_i}\right) \\ \text{subject to} & -\mathbf{P}^T\mathbf{w} \preceq \mathbf{c} \end{array},$$

with variable $\mathbf{w} \in \mathbb{R}^m$. Note that this is a GP in convex form with linear inequality constraints (i.e., monomial inequality constraints in the associated standard form GP).

4. (a) Introducing $\boldsymbol{\lambda}$ for the inequality constraint $\mathbf{x} \succeq \mathbf{0}$, and ν for the equality constraint $\mathbf{1}^T\mathbf{x} = 1$, it follows that the KKT conditions are as follows.

- *Primal feasibility:*
$$\mathbf{x}^\star \succeq \mathbf{0}\,,\ \mathbf{1}^T\mathbf{x}^\star = 1\,.$$
- *Dual feasibility:*
$$\boldsymbol{\lambda}^\star \succeq \mathbf{0}\,.$$
- *Complementary slackness:*
$$\lambda_k^\star\,(-x_k^\star) = 0\,,\ k = 1, \ldots, n\,.$$
- *Stationarity:*
$$\nabla f_0(\mathbf{x}^\star) - \boldsymbol{\lambda}^\star + \nu^\star \mathbf{1} = \mathbf{0}\,.$$

Here, $f_0(\mathbf{x}) = -\log\big(\mathbf{a}^T\mathbf{x}\big) - \log\big(\mathbf{b}^T\mathbf{x}\big)$, and so we have

$$\frac{\partial f_0}{\partial x_k} = -\frac{1}{\mathbf{a}^T\mathbf{x}}a_k - \frac{1}{\mathbf{b}^T\mathbf{x}}b_k \Longleftrightarrow \nabla f_0(\mathbf{x}) = -\frac{1}{\mathbf{a}^T\mathbf{x}}\mathbf{a} - \frac{1}{\mathbf{b}^T\mathbf{x}}\mathbf{b}\,.$$

Thus, the stationarity condition becomes

$$\boldsymbol{\lambda}^\star = \nu^\star\mathbf{1} - \frac{1}{\mathbf{a}^T\mathbf{x}^\star}\mathbf{a} - \frac{1}{\mathbf{b}^T\mathbf{x}^\star}\mathbf{b}\,.$$

Combining the results so far, we have the following simplifed KKT conditions.

$$\begin{aligned} &\mathbf{x}^\star \succeq \mathbf{0}\,,\ \mathbf{1}^T\mathbf{x}^\star = 1\,,\ \frac{1}{\mathbf{a}^T\mathbf{x}^\star} + \frac{1}{\mathbf{b}^T\mathbf{x}^\star}\mathbf{b} \preceq \nu^\star\mathbf{1}\,,\\ &x_k^\star\left(\nu^\star - \frac{1}{\mathbf{a}^T\mathbf{x}^\star}a_k - \frac{1}{\mathbf{b}^T\mathbf{x}^\star}b_k\right) = 0\,,\ k = 1, \ldots, n\,. \end{aligned} \tag{1}$$

We now show that $\mathbf{x}^\star = (1/2, 0, \ldots, 0, 1/2)$ and $\nu^\star = 2$ satisfy all of the conditions in (1) and are hence primal and dual optimal, respectively.
The feasibility conditions $\mathbf{x}^\star \succeq \mathbf{0}$, $\mathbf{1}^T\mathbf{x}^\star = 1$ obviously hold, and the complementary slackness conditions are trivially satisfied for $k = 2, \ldots, n-2$. Thus, from (1), it remains to verify the inequalities

$$\frac{a_k}{\mathbf{a}^T\mathbf{x}^\star} + \frac{b_k}{\mathbf{b}^T\mathbf{x}^\star} \leq \nu^\star\,,\ k = 1, \ldots, n\,, \tag{2}$$

and the complementary slackness conditions

$$x_k^\star\left(\nu^\star - \frac{1}{\mathbf{a}^T\mathbf{x}^\star}a_k - \frac{1}{\mathbf{b}^T\mathbf{x}^\star}b_k\right) = 0\,,\ k = 1, n\,. \tag{3}$$

For $\mathbf{x}^\star = (1/2, 0, \ldots, 0, 1/2)$ and $\nu^\star = 2$, the inequality (2) holds with equality for $k = 1$ and $k = n$, since we have

$$\frac{a_1}{\mathbf{a}^T\mathbf{x}^\star} + \frac{b_1}{\mathbf{b}^T\mathbf{x}^\star} = \frac{2a_1}{a_1 + a_n} + \frac{2/a_1}{1/a_1 + 1/a_n} = \frac{2a_1}{a_1 + a_n} + \frac{2a_n}{a_n + a_1} = 2\,,$$

for $k = 1$ and

$$\frac{a_n}{\mathbf{a}^T\mathbf{x}^\star} + \frac{b_n}{\mathbf{b}^T\mathbf{x}^\star} = \frac{2a_n}{a_1 + a_n} + \frac{2/a_n}{1/a_1 + 1/a_n} = \frac{2a_n}{a_1 + a_n} + \frac{2a_1}{a_n + a_1} = 2\,,$$

for $k = n$. Therefore also (3) is satisfied for $k = 1, n$ as desired. The remaining inequalities in (2) reduce to

$$\frac{a_k}{\mathbf{a}^T\mathbf{x}^\star} + \frac{b_k}{\mathbf{b}^T\mathbf{x}^\star} = \frac{2a_k}{a_1 + a_n} + \frac{2/a_k}{1/a_1 + 1/a_n} = 2\frac{a_k + a_1 a_n/a_k}{a_1 + a_n} \leq 2\,,\ \ k = 2, \ldots, n-1\,,$$

which is equivalent to

$$\frac{a_k + a_1 a_n/a_k}{a_1 + a_n} \leq 1\,,\ \ k = 2, \ldots, n-1\,. \tag{4}$$

To show that (4) is valid, note that the function $g(t) \triangleq \frac{t + a_1 a_n/t}{a_1 + a_n}$ is convex for $t \in \mathbb{R}_{++}$. Since the inequality in (4) holds with equality for $k = 1$ and $k = n$, it follows that

$$\frac{t + a_1 a_n/t}{a_1 + a_n} \leq 1\,,\ \ \forall\, t \in [a_n, a_1]\ .$$

Hence, $\mathbf{x}^\star = (1/2, 0, \ldots, 0, 1/2)$ and $\nu^\star$ satisfy all of the KKT conditions of (1) and are thus primal and dual optimal, respectively.

(b) Express $\mathbf{A}$ in terms of its eigenvalue decomposition as $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$, and define $a_k \triangleq \lambda_k$, $b_k \triangleq 1/\lambda_k$, and $x_k = \left[\mathbf{Q}^T\mathbf{u}\right]_k^2$. Note that we clearly have $x_k \geq 0$ for all k. Also, as $||\mathbf{u}||_2 = 1$, we have

$$\mathbf{1}^T\mathbf{x} = \sum_{k=1}^{n} \left[\mathbf{Q}^T\mathbf{u}\right]_k^2 = \left|\left|\mathbf{Q}^T\mathbf{u}\right|\right|_2^2 = \mathbf{u}^T\mathbf{Q}\mathbf{Q}^T\mathbf{u} = \mathbf{u}^T\mathbf{u} = 1\,.$$

Thus, the choice of $x_k = \left[\mathbf{Q}^T\mathbf{u}\right]_k^2$ is indeed feasible. Continuing further, note that

$$\mathbf{a}^T\mathbf{x} = \sum_{k=1}^{n} \lambda_k \left[\mathbf{Q}^T\mathbf{u}\right]_k^2 = \mathbf{u}^T\mathbf{A}\mathbf{u}\,,$$

and that

$$\mathbf{b}^T\mathbf{x} = \sum_{k=1}^{n} (1/\lambda_k) \left[\mathbf{Q}^T\mathbf{u}\right]_k^2 = \mathbf{u}^T\mathbf{A}^{-1}\mathbf{u}\,.$$

Hence, from the result proven in part (a), we have

$$-\log\!\left(\mathbf{u}^T\mathbf{A}\mathbf{u}\right) - \log\!\left(\mathbf{u}^T\mathbf{A}^{-1}\mathbf{u}\right) \geq -\log(\lambda_1/2 + \lambda_n/2) - \log(1/\left(2\lambda_1\right) + 1/\left(2\lambda_n\right))\ .$$

After some algebraic manipulation, this becomes

$$\left(\mathbf{u}^T\mathbf{A}\mathbf{u}\right)\left(\mathbf{u}^T\mathbf{A}^{-1}\mathbf{u}\right) \leq \left(\frac{\lambda_1}{2} + \frac{\lambda_n}{2}\right)\left(\frac{1}{2\lambda_1} + \frac{1}{2\lambda_n}\right) = \frac{1}{4}\left(\lambda_1 + \lambda_n\right)\left(\lambda_1^{-1} + \lambda_n^{-1}\right)\ .$$

This can be further simplified to

$$\left(\mathbf{u}^T\mathbf{A}\mathbf{u}\right)\left(\mathbf{u}^T\mathbf{A}^{-1}\mathbf{u}\right) \leq \frac{1}{4}\left(\lambda_1/\lambda_n + 2 + \frac{\lambda_n}{\lambda_1}\right) = \frac{1}{4}\left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\right)^2\ .$$

Taking square roots of both sides of the above inequality yields

$$2\left(\mathbf{u}^T\mathbf{A}\mathbf{u}\right)^{1/2}\left(\mathbf{u}^T\mathbf{A}^{-1}\mathbf{u}\right)^{1/2} \leq \sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\,,$$

for all $\mathbf{u}$ with $||\mathbf{u}||_2 = 1$, which is Kantorovich's inequality.

***5.** Define the following quantities.

$$\mathbf{A} \triangleq \begin{bmatrix} -2\mathbf{y}_1^T & 1 \\ \vdots & \vdots \\ -2\mathbf{y}_m^T & 1 \end{bmatrix}, \mathbf{b} \triangleq \begin{bmatrix} d_1^2 - ||\mathbf{y}_1||_2^2 \\ \vdots \\ d_m^2 - ||\mathbf{y}_m||_2^2 \end{bmatrix}, \mathbf{C} \triangleq \begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{n\times 1} \\ \mathbf{0}_{1\times n} & 0 \end{bmatrix}, \mathbf{f} \triangleq \begin{bmatrix} \mathbf{0}_{n\times 1} \\ -1/2 \end{bmatrix}.$$

Also, define $\mathbf{z} \triangleq (\mathbf{x}, t)$. Then, with this notation, the problem becomes

$$\begin{array}{ll} \text{minimize} & ||\mathbf{A}\mathbf{x} - \mathbf{b}||_2^2 \\ \text{subject to} & \mathbf{z}^T\mathbf{C}\mathbf{z} + 2\mathbf{f}^T\mathbf{z} = 0 \end{array}.$$

Introducing ν for the equality constraint, we obtain the following for the Lagrangian.

$$\begin{aligned} L(\mathbf{z}, \nu) &= ||\mathbf{A}\mathbf{z} - \mathbf{b}||_2^2 + \nu\left(\mathbf{z}^T\mathbf{C}\mathbf{z} + 2\mathbf{f}^T\mathbf{z}\right), \\ &= \mathbf{z}^T\mathbf{A}^T\mathbf{A}\mathbf{z} - 2\mathbf{b}^T\mathbf{A}\mathbf{z} + \mathbf{b}^T\mathbf{b} + \mathbf{z}^T(\nu\mathbf{C})\mathbf{z} + 2\nu\mathbf{f}^T\mathbf{z}, \\ &= \mathbf{z}^T\left(\mathbf{A}^T\mathbf{A} + \nu\mathbf{C}\right)\mathbf{z} - 2\left(\mathbf{A}^T\mathbf{b} - \nu\mathbf{f}\right)^T\mathbf{z} + ||\mathbf{b}||_2^2. \end{aligned}$$

Note that this is bounded below as a function of $\mathbf{z}$ if and only if

$$\mathbf{A}^T\mathbf{A} + \nu\mathbf{C} \succeq \mathbf{0},\ \mathbf{A}^T\mathbf{b} - \nu\mathbf{f} \in \mathcal{R}\left(\mathbf{A}^T\mathbf{A} + \nu\mathbf{C}\right).$$

Therefore, the KKT conditions are as follows.

- *Primal feasibility:*

$$\mathbf{z}^T\mathbf{C}\mathbf{z} + 2\mathbf{f}^T\mathbf{z} = 0.$$

- *Dual feasibility:*

$$\mathbf{A}^T\mathbf{A} + \nu\mathbf{C} \succeq \mathbf{0},\ \mathbf{A}^T\mathbf{b} - \nu\mathbf{f} \in \mathcal{R}\left(\mathbf{A}^T\mathbf{A} + \nu\mathbf{C}\right).$$

- *Stationarity:*

$$\left(\mathbf{A}^T\mathbf{A} + \nu\mathbf{C}\right)\mathbf{z} = \mathbf{A}^T\mathbf{b} - \nu\mathbf{f}.$$

(Note that this implies the range condition for dual feasibility.)

Method 1: We derive the dual problem. If ν is feasible, then the dual function is given by

$$g(\nu) = -\left(\mathbf{A}^T\mathbf{b} - \nu\mathbf{f}\right)^T\left(\mathbf{A}^T\mathbf{A} + \nu\mathbf{C}\right)^{\#}\left(\mathbf{A}^T\mathbf{b} - \nu\mathbf{f}\right) + ||\mathbf{b}||_2^2.$$

So, the dual problem can be expressed as the SDP

$$\begin{array}{ll} \text{maximize} & -s + ||\mathbf{b}||_2^2 \\ \text{subject to} & \begin{bmatrix} \mathbf{A}^T\mathbf{A} + \nu\mathbf{C} & \mathbf{A}^T\mathbf{b} - \nu\mathbf{f} \\ \left(\mathbf{A}^T\mathbf{b} - \nu\mathbf{f}\right)^T & s \end{bmatrix} \succeq \mathbf{0} \end{array},$$

which is equivalent to the following SDP

$$\begin{array}{ll} \text{minimize} & s - ||\mathbf{b}||_2^2 \\ \text{subject to} & \begin{bmatrix} \mathbf{A}^T\mathbf{A} + \nu\mathbf{C} & \mathbf{A}^T\mathbf{b} - \nu\mathbf{f} \\ \left(\mathbf{A}^T\mathbf{b} - \nu\mathbf{f}\right)^T & s \end{bmatrix} \succeq \mathbf{0} \end{array},$$

Solving this in `cvx` gives $\nu^\star = 0.5898$. Using $\nu^\star$ in the stationarity condition, we get

$$\mathbf{z}^\star = \left(\mathbf{A}^T\mathbf{A} + \nu^\star\mathbf{C}\right)^{-1}\left(\mathbf{A}^T\mathbf{b} - \nu^\star\mathbf{f}\right) = (1.3269, 0.6446, 2.1765)\ .$$

Hence, $\mathbf{x}^\star = (1.3269, 0.6446)$.

Method 2: Alternatively, we can solve the KKT conditions directly. To simplify the equations, we make a change of variables

$$\mathbf{w} = \mathbf{Q}^T\mathbf{L}^T\mathbf{z}\,,$$

where $\mathbf{L}$ is the lower triangular matrix obtained from the Cholesky decomposition of $\mathbf{A}^T\mathbf{A}$, i.e., $\mathbf{A}^T\mathbf{A} = \mathbf{L}\mathbf{L}^T$, and $\mathbf{Q}$ is the matrix of eigenvectors of $\mathbf{L}^{-1}\mathbf{C}\mathbf{L}^{-T}$, i.e., $\mathbf{L}^{-1}\mathbf{C}\mathbf{L}^{-T} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$. This transforms the KKT conditions to

$$\mathbf{w}^T\mathbf{\Lambda}\mathbf{w} + 2\mathbf{g}^T\mathbf{w} = 0\,,\ \mathbf{I} + \nu\mathbf{\Lambda} \succeq \mathbf{0}\,,\ \left(\mathbf{I} + \nu\mathbf{\Lambda}\right)\mathbf{w} = \mathbf{h} - \nu\mathbf{g}\,,$$

where we have

$$\mathbf{g} = \mathbf{Q}^T\mathbf{L}^{-1}\mathbf{f}\,,\ \mathbf{h} = \mathbf{Q}^T\mathbf{L}^{-1}\mathbf{A}^T\mathbf{b}\,.$$

From the last equation of the KKT conditions, we find

$$w_k = \frac{h_k - \nu g_k}{1 + \nu\lambda_k}\,,\ k = 1, \ldots, n+1, .$$

Substituting this into the first equation of the KKT conditions, we get the following nonlinear equation in ν.

$$r(\nu) = \sum_{k=1}^{n+1}\left(\frac{\lambda_k\left(h_k - \nu g_k\right)^2}{\left(1 + \nu\lambda_k\right)^2} + \frac{2g_k\left(h_k - \nu g_k\right)}{1 + \nu\lambda_k}\right) = 0\,.$$

In our example, the eigenvalues are

$$\lambda_1 = 0.5104\,,\ \lambda_2 = 0.2735\,,\ \lambda_3 = 0\,.$$

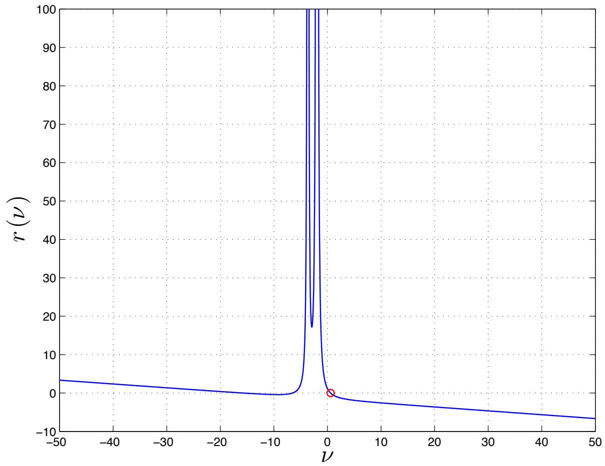
Plots of the function $r(\nu)$ for this example are shown in Figures 4(a) and (b). In Figure 4(a), we have a zoomed out plot showing all three solutions to $r(\nu) = 0$, whereas in Figure 4(b), we have a zoomed out plot showing the correct solution.

The correct solution of $r(\nu) = 0$ is the one that satisfies $1 + \nu\lambda_k \geq 0$ for $k = 1, \ldots, n+1$, i.e., the solution to the right of the two singularities in this case. This solution can be determined by using Newton's method by repeating the iteration

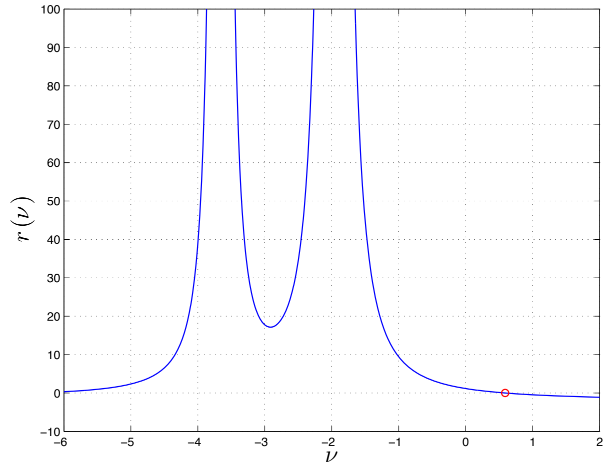
$$\nu := \nu - \frac{r(\nu)}{r'(\nu)}$$

a few times, starting at a value close to the desired solution. This gives $\nu^\star = 0.5896$, which is very close to the value obtained from `cvx`. From $\nu^\star$, we determine $\mathbf{x}^\star$ as in the first method.

A contour plot of the objective f_0 for the given problem data, along with the sensor position vectors $\mathbf{y}_k$ and optimal source position vector $\mathbf{x}^\star$ is shown in Figure 5.



(a)



(b)

Figure 4: Plots of the nonlinear function $r(\nu)$ used to determine the optimal ν which satisfies the KKT conditions: (a) zoomed out plot showing all three solutions to $r(\nu) = 0$ and (b) zoomed in plot showing a close-up of the correct solution to $r(\nu) = 0$.

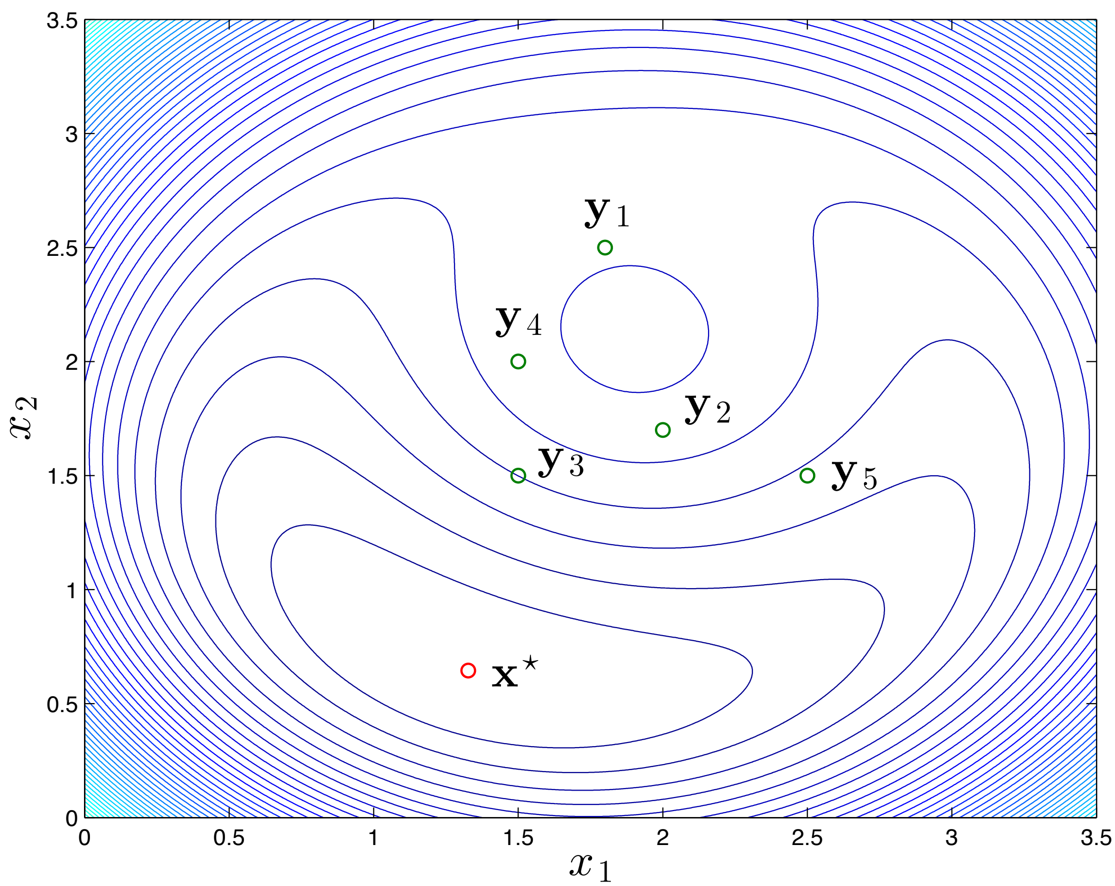


Figure 5: Contour plot of the objective $f_0(x_1, x_2)$ for the given problem data, with the sensor position vectors $\mathbf{y}_k$ and optimal source position vector $\mathbf{x}^\star$ indicated by circles.