5.1 *A simple example.* Consider the optimization problem

$$\begin{array}{ll} \text{minimize} & x^2+1 \\ \text{subject to} & (x-2)(x-4) \leq 0, \end{array}$$

with variable $x \in \mathbf{R}$.

(a) *Analysis of primal problem.* Give the feasible set, the optimal value, and the optimal solution.

(b) *Lagrangian and dual function.* Plot the objective x^2+1 versus x. On the same plot, show the feasible set, optimal point and value, and plot the Lagrangian $L(x,\lambda)$ versus x for a few positive values of λ. Verify the lower bound property ($p^\star \geq \inf_x L(x,\lambda)$ for $\lambda \geq 0$). Derive and sketch the Lagrange dual function g.

(c) *Lagrange dual problem.* State the dual problem, and verify that it is a concave maximization problem. Find the dual optimal value and dual optimal solution $\lambda^\star$. Does strong duality hold?

(d) *Sensitivity analysis.* Let $p^\star(u)$ denote the optimal value of the problem

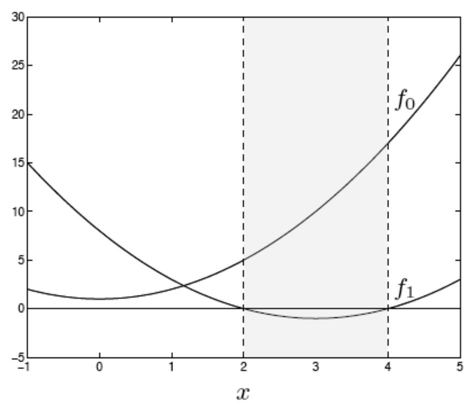
$$\begin{array}{ll} \text{minimize} & x^2+1 \\ \text{subject to} & (x-2)(x-4) \leq u, \end{array}$$

as a function of the parameter u. Plot $p^\star(u)$. Verify that $dp^\star(0)/du = -\lambda^\star$.

Solution.

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

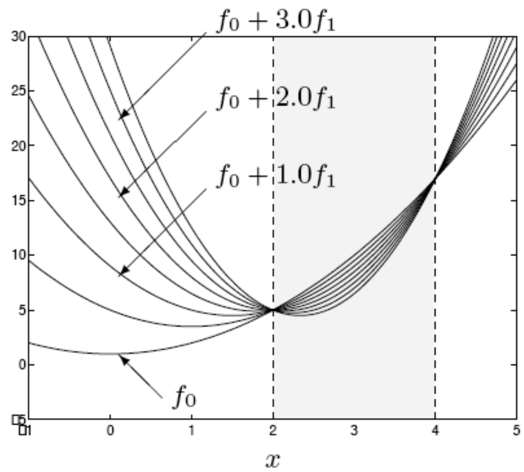
The plot shows f_0 and f_1.



(b) The Lagrangian is

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

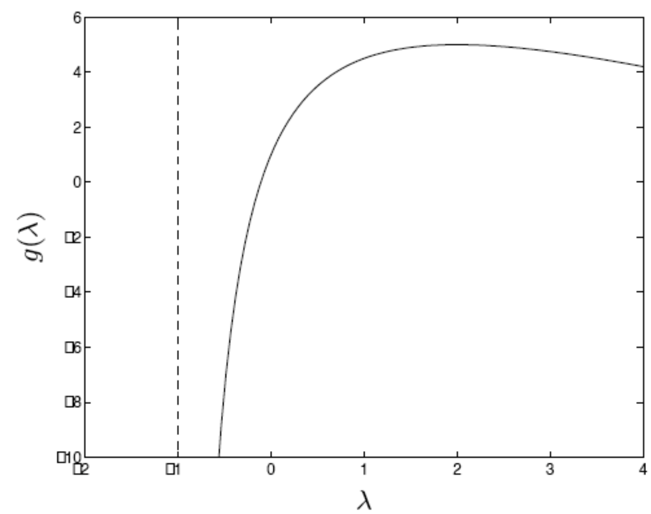
The plot shows the Lagrangian $L(x,\lambda) = f_0 + \lambda f_1$ as a function of x for different values of $\lambda \geq 0$. Note that the minimum value of $L(x,\lambda)$ over x (*i.e.*, $g(\lambda)$) is always less than $p^\star$. It increases as λ varies from 0 toward 2, reaches its maximum at $\lambda = 2$, and then decreases again as λ increases above 2. We have equality $p^\star = g(\lambda)$ for $\lambda = 2$.



For $\lambda > -1$, the Lagrangian reaches its minimum at $\tilde{x} = 3\lambda/(1+\lambda)$. For $\lambda \le -1$ it is unbounded below. Thus

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

which is plotted below.



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.

(c) The Lagrange dual problem is

$$\begin{array}{ll} \text{maximize} & -9\lambda^2/(1+\lambda) + 1 + 8\lambda \\ \text{subject to} & \lambda \ge 0. \end{array}$$

The dual optimum occurs at $\lambda = 2$, with $d^\star = 5$. So for this example we can directly observe that strong duality holds (as it must — Slater's constraint qualification is satisfied).

(d) The perturbed problem is infeasible for $u < -1$, since $\inf_x(x^2 - 6x + 8) = -1$. For $u \ge -1$, the feasible set is the interval

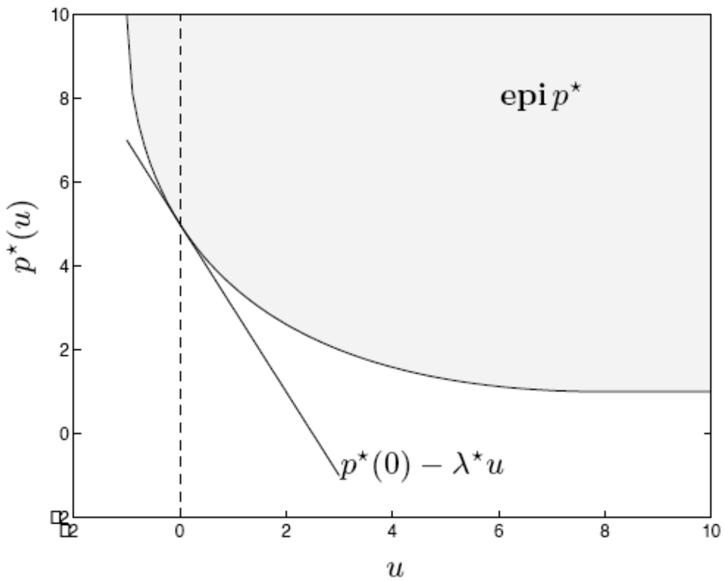
$$[3 - \sqrt{1+u}, 3 + \sqrt{1+u}],$$

given by the two roots of $x^2 - 6x + 8 = u$. For $-1 \le u \le 8$ the optimum is $x^\star(u) = 3 - \sqrt{1+u}$. For $u \ge 8$, the optimum is the unconstrained minimum of f_0,

$x^\star(u) = 3 - \sqrt{1+u}$. For $u \geq 8$, the optimum is the unconstrained minimum of f_0, *i.e.*, $x^\star(u) = 0$. In summary,

$$p^\star(u) = \begin{cases} \infty & u < -1 \\ 11 + u - 6\sqrt{1+u} & -1 \leq u \leq 8 \\ 1 & u \geq 8. \end{cases}$$

The figure shows the optimal value function $p^\star(u)$ and its epigraph.



Finally, we note that $p^\star(u)$ is a differentiable function of u, and that

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

5.20 *Dual of channel capacity problem.* Derive a dual for the problem

$$\begin{array}{ll} \text{minimize} & -c^T x + \sum_{i=1}^m y_i \log y_i \\ \text{subject to} & Px = y \\ & x \succeq 0, \quad \mathbf{1}^T x = 1, \end{array}$$

where $P \in \mathbf{R}^{m\times n}$ has nonnegative elements, and its columns add up to one (*i.e.*, $P^T\mathbf{1} = \mathbf{1}$). The variables are $x \in \mathbf{R}^n$, $y \in \mathbf{R}^m$. (For $c_j = \sum_{i=1}^m p_{ij}\log p_{ij}$, the optimal value is, up to a factor $\log 2$, the negative of the capacity of a discrete memoryless channel with channel transition probability matrix P; see exercise 4.57.)

Simplify the dual problem as much as possible.

Solution. The Lagrangian is

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

The minimum over x is bounded below if and only if

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

To minimize over y, we set the derivative with respect to y_i equal to zero, which gives $\log y_i + 1 - z_i = 0$, and conclude that

$$\inf_{y_i \geq 0}(y_i \log y_i - z_i y_i) = -e^{z_i - 1}.$$

The dual function is

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

The dual problem is

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

This can be simplified by introducing a variable $w = z + \nu\mathbf{1}$ (and using the fact that $\mathbf{1} = P^T\mathbf{1}$), which gives

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

Finally we can easily maximize the objective function over ν by setting the derivative equal to zero (the optimal value is $\nu = -\log(\sum_i e^{1-w_i})$), which leads to

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

This is a geometric program, in convex form, with linear inequality constraints (*i.e.*, monomial inequality constraints in the associated geometric program).

5.26 Consider the QCQP

$$\begin{array}{ll} \text{minimize} & x_1^2 + x_2^2 \\ \text{subject to} & (x_1-1)^2 + (x_2-1)^2 \le 1 \\ & (x_1-1)^2 + (x_2+1)^2 \le 1 \end{array}$$

with variable $x \in \mathbf{R}^2$.

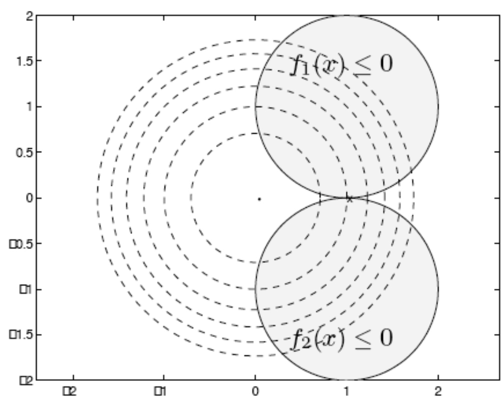
(a) Sketch the feasible set and level sets of the objective. Find the optimal point $x^\star$ and optimal value $p^\star$.

(b) Give the KKT conditions. Do there exist Lagrange multipliers $\lambda_1^\star$ and $\lambda_2^\star$ that prove that $x^\star$ is optimal?

(c) Derive and solve the Lagrange dual problem. Does strong duality hold?

Solution.

(a) The figure shows the feasible set (the intersection of the two shaded disks) and some contour lines of the objective function. There is only one feasible point, $(1, 0)$, so it is optimal for the primal problem, and we have $p^\star = 1$.



(b) The KKT conditions are

$$\begin{gathered} (x_1-1)^2 + (x_2-1)^2 \le 1, \quad (x_1-1)^2 + (x_2+1)^2 \le 1, \\ \lambda_1 \ge 0, \qquad \lambda_2 \ge 0 \\ 2x_1 + 2\lambda_1(x_1-1) + 2\lambda_2(x_1-1) = 0 \\ 2x_2 + 2\lambda_1(x_2-1) + 2\lambda_2(x_2+1) = 0 \\ \lambda_1((x_1-1)^2 + (x_2-1)^2 - 1) = \lambda_2((x_1-1)^2 + (x_2+1)^2 - 1) = 0. \end{gathered}$$

At $x = (1, 0)$, these conditions reduce to

$$\lambda_1 \ge 0, \qquad \lambda_2 \ge 0, \qquad 2 = 0, \qquad -2\lambda_1 + 2\lambda_2 = 0,$$

which (clearly, in view of the third equation) have no solution.

(c) The Lagrange dual function is given by

$$g(\lambda_1, \lambda_2) = \inf_{x_1, x_2} L(x_1, x_2, \lambda_1, \lambda_2)$$

where

$$\begin{aligned} & L(x_1, x_2, \lambda_1, \lambda_2) \\ &= x_1^2 + x_2^2 + \lambda_1((x_1-1)^2 + (x_2-1)^2 - 1) + \lambda_2((x_1-1)^2 + (x_2+1)^2 - 1). \end{aligned}$$

L reaches its minimum for

$$x_1 = \frac{\lambda_1 + \lambda_2}{1 + \lambda_1 + \lambda_2}, \qquad x_2 = \frac{-\lambda_1 + \lambda_2}{1 + \lambda_1 + \lambda_2},$$

and we find

$$g(\lambda_1, \lambda_2) \quad = \quad \begin{cases} -\frac{(\lambda_1+\lambda_2)^2+(\lambda_1-\lambda_2)^2}{1+\lambda_1+\lambda_2} - \lambda_1 - \lambda_2 & 1 + \lambda_1 + \lambda_2 \geq 0 \\ -\infty & \text{otherwise,} \end{cases}$$

where we interpret $a/0 = 0$ if $a = 0$ and as $-\infty$ if $a < 0$. The Lagrange dual problem is given by

$$\begin{array}{ll} \text{maximize} & -((\lambda_1 - \lambda_2)^2 - \lambda_1 - \lambda_2)/(1 + \lambda_1 + \lambda_2) \\ \text{subject to} & \lambda_1, \lambda_2 \geq 0. \end{array}$$

Since g is symmetric, the optimum (if it exists) occurs with $\lambda_1 = \lambda_2$. The dual function then simplifies to

$$g(\lambda_1, \lambda_1) = \frac{2\lambda_1}{2\lambda_1 + 1}.$$

We see that $g(\lambda_1, \lambda_2)$ tends to 1 as $\lambda_1 \to \infty$. We have $d^\star = p^\star = 1$, but the dual optimum is not attained.

Recall that the KKT conditions only hold if (1) strong duality holds, (2) the primal optimum is attained, and (3) the dual optimum is attained. In this example, the KKT conditions fail because the dual optimum is not attained.