Homework 3
Convex and Nonsmooth Optimization

Evan Toler

Feb. 25, 2020

1. Consider the primal problem of minimizing $f_0(x) = \frac{1}{2}x^T Qx$ subject to $h(x) = Ax - b = 0$, where Q is a symmetric positive definite $n \times n$ matrix and A is $m \times n$, with $m \leq n$, with full rank m (in other words, A has m linearly independent rows). There are no inequality constraints.

   (a) Write down the Lagrangian $L(x, \nu)$.

   Solution:
   $$L(x, \nu) = \frac{1}{2}x^T Qx + \nu^T(Ax - b)$$

   (b) Since $L(x, \nu)$ is convex, differentiable and bounded below in x, set its gradient to zero to find its minimizer and write down a formula for the Lagrange dual function $g(\nu) = \inf_x L(x, \nu)$ (as inf can be replaced by min, in this case).

   Solution: Differentiating yields
   $$\nabla_x L(x, \nu) = Qx + A^T \nu \overset{\text{set}}{=} 0$$
   which gives an expression for the minimizing x as
   $$x = -Q^{-1}A^T\nu.$$
   Next, we establish some intermediary facts and observations. First, since $Q \succ 0$, its inverse is positive definite as well. Moreover, since A has full row rank, the matrix $AQ^{-1}A^T$ is positive definite and hence invertible. We make use of this fact in the next part of the question. For now, we may write the dual objective as
   $$\begin{aligned} g(\nu) &= L(-Q^{-1}A^T\nu, \nu) \\ &= \frac{1}{2}\nu^T AQ^{-1}QQ^{-1}A^T\nu + \nu^T(-AQ^{-1}A^T\nu - b) \\ &= -\frac{1}{2}\nu^T AQ^{-1}A^T\nu - \nu^T b. \end{aligned}$$

   (c) Find the maximizer ν^* of the Lagrange dual function $g(\nu)$ (which is concave) by setting its gradient to zero. What is the dual optimal value $d^* = g(\nu^*)$?

   Solution: Differentiating g gives
   $$\nabla_\nu g = -AQ^{-1}A^T\nu - b \overset{\text{set}}{=} 0$$

and so the maximizing value for ν is

$$\nu^* = -(AQ^{-1}A^T)^{-1}b.$$

Note that the system matrix is invertible since it is positive definite. The corresponding dual objective value is

$$\begin{aligned} d^* = g(\nu^*) &= -\frac{1}{2}[b^T(AQ^{-1}A^T)^{-1}]AQ^{-1}A^T[(AQ^{-1}A^T)^{-1}b] + b^T(AQ^{-1}A^T)^{-1}b \\ &= \frac{1}{2}b^T(AQ^{-1}A^T)^{-1}b. \end{aligned}$$

(d) Find the associated $\widehat{x}$ attaining the minimizer of the Lagrangian $L(x, \nu^*)$.

Solution: The associated $\widehat{x}$ is given from part (b) as

$$\begin{aligned} \widehat{x} &= -Q^{-1}A^T\nu^* \\ &= Q^{-1}A^T(AQ^{-1}A^T)^{-1}b. \end{aligned}$$

(e) Check whether $\widehat{x}$ is feasible for the primal problem (whether it satisfies $Ax = b$).

Solution: We can compute directly that

$$\begin{aligned} A\widehat{x} &= AQ^{-1}A^T(AQ^{-1}A^T)^{-1}b \\ &= b. \end{aligned}$$

Therefore the point $\widehat{x}$ is primal feasible.

(f) Find the primal value $f_0(\widehat{x})$. If $\widehat{x}$ is primal feasible, then the optimal primal value $p^* \leq f_0(\widehat{x})$.

Solution:

$$\begin{aligned} f_0(\widehat{x}) &= \frac{1}{2}\widehat{x}^TQ\widehat{x} \\ &= \frac{1}{2}[b^T(AQ^{-1}A^T)^{-1}AQ^{-1}]Q[Q^{-1}A^T(AQ^{-1}A^T)^{-1}b] \\ &= \frac{1}{2}b^T(AQ^{-1}A^T)^{-1}AQ^{-1}A^T(AQ^{-1}A^T)^{-1}b \\ &= \frac{1}{2}b^T(AQ^{-1}A^T)^{-1}b. \end{aligned}$$

Since $\widehat{x}$ is primal feasible, $p^* \leq f_0(\widehat{x})$.

(g) Do you conclude that there is no duality gap, i.e., that $d^* = p^*$?

Solution: Since $f_0(\widehat{x}) = d^*$ and $p^* \leq f_0(\widehat{x})$, we conclude that $p^* \leq d^*$. On the other hand, weak duality gives that $d^* \leq p^*$. Hence there must be no duality gap, i.e., $d^* = p^*$.

Note: We did exactly this computation in class in the special case $Q = I$.

2. (BV 5.1) 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 $x \in \mathbb{R}$.

(a) Give the feasible set, the optimal value, and the optimal solution.

Solution: The feasible set is all $x \in [2,4]$. By inspection, the optimal solution is the element of the feasible set with least absolute value. That is, $x^* = 2$, which gives an optimal value of $p^* = 5$.

(b) 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 ($p^* \geq \inf_x L(x,\lambda)$ for $\lambda \geq 0$). Derive and sketch the Lagrange dual function g.

Solution: The Lagrangian is given by

$$L(x,\lambda) = x^2 + 1 + \lambda(x-2)(x-4).$$

In Figure 1, we illustrate the desired properties of the Lagrangian. In particular, we observe for $\lambda = 2$, the lower bound property holds with equality, and for $\lambda \neq 2$, it holds strictly.

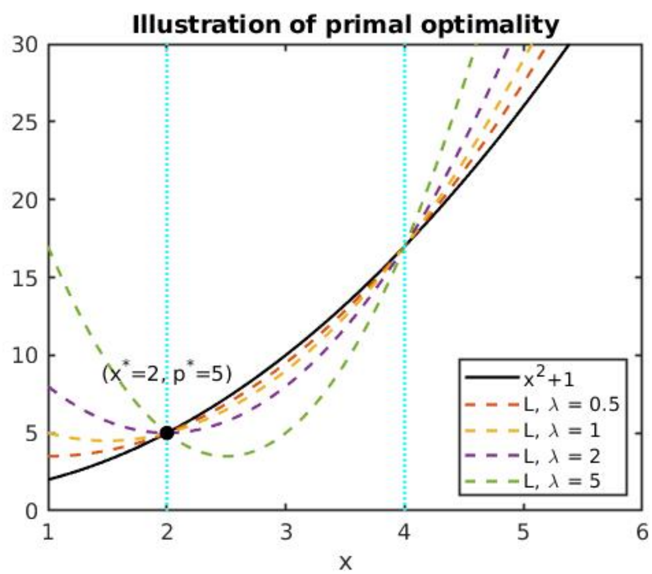

Figure 1: Objective function and Lagrangian for different $\lambda \geq 0$. The dotted lines indicate the boundary of the feasible set. In each case, $p^* \geq \inf_x L(x,\lambda)$.

We derive the Lagrange dual function

$$\begin{aligned} g(\lambda) &= \inf_x x^2 + 1 + \lambda(x-2)(x-4) \\ &= \inf_x (1+\lambda)x^2 - 6\lambda x + (8\lambda+1). \end{aligned}$$

Since $L(x,\lambda)$ is convex for $\lambda > -1$, we may differentiate to find the infimum.

$$2(1+\lambda)x - 6\lambda = 0 \quad \implies \quad x = \frac{3\lambda}{1+\lambda}.$$

We conclude that the dual function for $\lambda > -1$ is given by

$$\begin{aligned} g(\lambda) &= L(3\lambda/(1+\lambda), \lambda) \\ &= (1+\lambda)(\frac{3\lambda}{1+\lambda})^2 - 6\lambda\frac{3\lambda}{1+\lambda} + 8\lambda + 1 \\ &= \frac{-9\lambda^2}{1+\lambda} + 8\lambda + 1. \end{aligned}$$

If $\lambda \leq -1$, then L is concave, and hence the infimum defining g is $-\infty$. So, the full characterization of the Lagrange dual function is

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

plotted below.

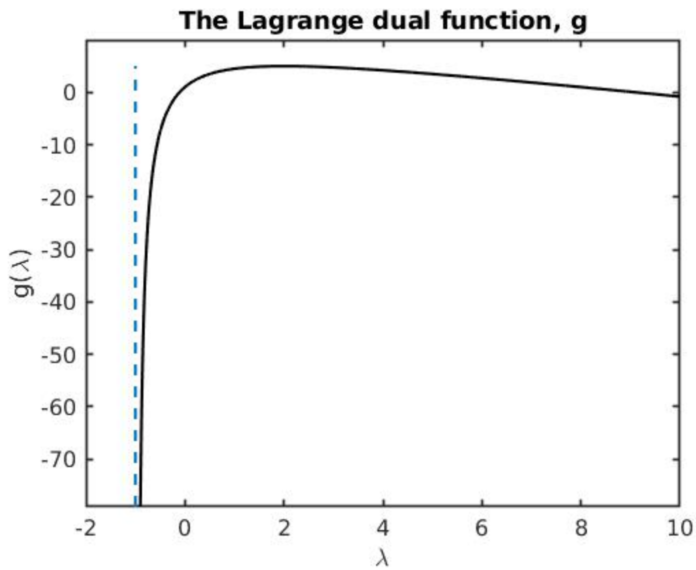

Figure 2: The Lagrange dual function for $\lambda > -1$, the domain of g

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

Solution: The Lagrange dual problem is

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

Writing $g_0(\lambda) = -9\lambda^2/(1+\lambda) + 8\lambda + 1$ as the objective function, we find after some straightforward calculus that

$$\frac{d^2}{d\lambda^2}(-g_0) = \frac{18}{(1+\lambda)^3},$$

which is strictly positive for all $\lambda \geq 0$. Hence we conclude that $-g_0$ is convex, and so g_0 is concave. Clearly, the feasible set is also convex, so the dual problem is indeed a concave maximization problem.

We may differentiate g_0 to find the optimal solution:

$$\frac{dg_0}{d\lambda} = -\frac{9\lambda^2 + 18\lambda}{(1+\lambda)^2} + 8 \overset{\text{set}}{=} 0.$$

Solving, we find either $\lambda = 2$ or $\lambda = -4$. Only the first candidate is feasible, and so $\lambda^* = 2$. The optimal dual value is

$$d^* = g(2) = 5 = p^*.$$

So, strong duality holds for this problem.

(d) Let $p^*(u)$ denote the optimal value of the problem

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

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

Solution: The feasible set for the modified problem is all $x \in [3 - \sqrt{1+u}, 3 + \sqrt{1+u}]$ for $u \ge -1$, and the problem is infeasible for $u < -1$. Since the global minimizer over $\mathbb{R}$ of $x^2 + 1$ is $x^* = 0$, we have that

$$p^*(u) = \begin{cases} (3 - \sqrt{1+u})^2 + 1, & -1 \le u < 8 \\ 1, & u \ge 8 \\ \infty, & \text{otherwise.} \end{cases}$$

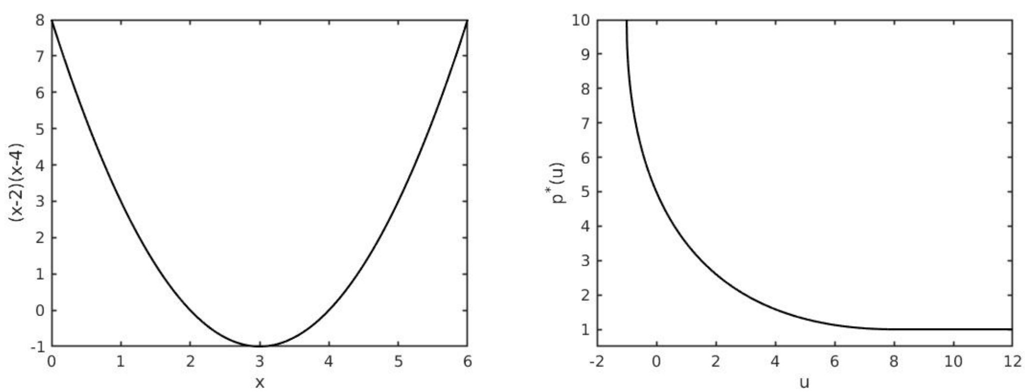

Figure 3: The original constraint $f_1(x) = (x-2)(x-4)$ and the optimal value $p^*(u)$

The modified Lagrangian is $L(x, \lambda; u) = x^2 + 1 + \lambda((x-2)(x-4) - u)$. The x which minimizes L for fixed λ is unchanged, and the new dual function $\widehat{g}$ is given by $\widehat{g}(\lambda) = g(\lambda) - \lambda u$. Therefore at $u = 0$, $\lambda^* = 2$ is unchanged. *Remark: Without retracing the previous parts of the problem, one could conclude that λ^* is unchanged since at $u = 0$, the problem is identical to the one in (a).*

Finally, we verify that

$$
\begin{aligned}
\frac{dp^*}{du}(0) &= \frac{d}{du}((3-\sqrt{1+u})^2+1)\bigg|_{u=0} \\
&= -\frac{3-\sqrt{1+u}}{\sqrt{1+u}}\bigg|_{u=0} \\
&= -2 \\
&= -\lambda^*.
\end{aligned}
$$

3. (BV 5.21) *A convex problem in which strong duality fails.* Consider the optimization problem

$$
\begin{array}{ll}
\text{minimize} & e^{-x} \\
\text{subject to} & x^2/y \le 0,
\end{array}
$$

with variables x and y, and domain $\mathcal{D} = \{(x,y) \mid y > 0\}$.

(a) Verify that this is a convex optimization problem. Find the optimal value.

Solution: The objective e^{-x} is convex since its epigraph is a convex set. The domain $\mathcal{D}$ is an open half plane, which is convex. The function $f_1(x,y) = x^2/y$ has Hessian

$$
H := \nabla^2 f_1 = \begin{pmatrix} 2/y & -2x/y^2 \\ -2x/y^2 & 2x^2y^3 \end{pmatrix}.
$$

We claim that for all $x, y \in \mathcal{D}$, H is positive semidefinite. Indeed, for arbitrary $v = (v_1, v_2) \in \mathbb{R}^2$,

$$
\begin{aligned}
v^T H v = \frac{2v_1^2}{y} - \frac{4xv_1v_2}{y^2} + \frac{2x^2v_2^2}{y^3} &\ge 0 \\
\iff v_1^2y^2 - 2xyv_1v_2 + x^2v_2^2 &\ge 0 \\
\iff (yv_1 - xv_2)^2 &\ge 0.
\end{aligned}
$$

The square of any scalar quantity is nonnegative, so our claim is true. H is positive semidefinite for all $x, y \in \mathcal{D}$ and so f_1 is convex. Since the domain, objective, and constraint are all convex, we conclude that this is a convex optimization problem.
Since y is strictly positive, the only feasible value of x is $x = 0$, and so we deduce that $p^* = 1$.

(b) Give the Lagrange dual problem, and find the optimal solution λ^* and optimal value d^* of the dual problem. What is the optimal duality gap?

Solution: The Lagrangian is $L(x, y, \lambda) = e^{-x} + \lambda x^2/y$. If $\lambda < 0$, then $g(\lambda) = \inf_{(x,y)\in\mathcal{D}} L = -\infty$ by taking $x \to \infty$.
If $\lambda \ge 0$, then $g(\lambda) = 0$ by taking $x \to \infty$ with $y \sim x^4$. So, the dual problem is

$$
\begin{array}{ll}
\text{maximize} & 0 \\
\text{subject to} & \lambda \ge 0.
\end{array}
$$

Any $\lambda^* \ge 0$ is optimal, and $d^* = 0$. The duality gap is $p^* - d^* = 1$.

(c) Does Slater's condition hold for this problem?

Solution: Slater's condition cannot hold for this problem. If it did, this would contradict the strong duality theorem—that convex problems satisfying Slater's condition have $p^* = d^*$.

Concretely, we saw in part (a) that the only feasible value of x is $x = 0$, which means the inequality constraint must hold with equality. Hence there can be no $(x, y) \in \text{relint } \mathcal{D}$ with $f_1(x, y) < 0$.

(d) What is the optimal value $p^*(u)$ of the perturbed problem

$$\begin{array}{ll} \text{minimize} & e^{-x} \\ \text{subject to} & x^2/y \leq u \end{array}$$

as a function of u? Verify that the global sensitivity inequality

$$p^*(u) \geq p^*(0) - \lambda^* u$$

does not hold.

Solution: If $u < 0$, then clearly the problem is infeasible and $p^*(u) = \infty$. If $u > 0$, then we may rewrite the constraint as $x \in [-\sqrt{uy}, \sqrt{uy}]$. Hence x and y can be made arbitrarily large, and so $p^*(u) = 0$. To summarize,

$$p^*(u) = \begin{cases} \infty, & u < 0 \\ 1, & u = 0 \\ 0, & u > 0. \end{cases}$$

The global sensitivity inequality, which would hold if the problem had strong duality, fails in this case. To illustrate, we may take $\lambda^* = 0$ as the optimal dual variable, from which we obtain the patently *untrue* inequality

$$p^*(u) \geq p^*(0).$$

In fact, the inequality holds in reverse! We have that $p^*(0) \geq p^*(u)$ for all $u \geq 0$.

4. Show that for convex problems of the form (5.25), that is (5.1) with $f_0, f_1, \ldots, f_m$ convex and $h_1, \ldots, h_p$ affine, the set $\mathcal{A}$ defined in BV (5.37) is convex. This is crucial to the proof of strong duality, because the separating hyperplane theorem is applied to separate $\mathcal{A}$ and $\mathcal{B}$.

Solution: Recall that the set $\mathcal{A}$ is given by

$$\mathcal{A} = \{(u, v, t) \in \mathbb{R}^{m+p+1} \mid \exists x \in \mathcal{D} \text{ such that } f(x) \leq u, h(x) = v, f_0(x) \leq t\}.$$

Here, we have written $f(x) \leq u$ to mean $f_i(x) \leq u_i$ for all $1 \leq i \leq m$, and similarly for $h(x) = v$.

To show that $\mathcal{A}$ is convex, consider two points $(u, t, v), (u', v', t') \in \mathcal{A}$. There exist x, x' correspondingly which thus satisfy

$$\begin{array}{lll} f(x) \leq u, & h(x) = v, & f_0(x) \leq t, \\ f(x') \leq u', & h(x') = v', & f_0(x') \leq t'. \end{array}$$

Moreover, since the problem is assumed to be convex, $\mathcal{D}$ is a convex set and hence

$$y := \theta x + (1 - \theta)x' \in \mathcal{D}.$$

By the convexity of f_i for all i,

$$f(y) \leq \theta f(x) + (1 - \theta) f(x') \leq \theta u + (1 - \theta) u'.$$

And by an identical argument, $f_0(y) \leq \theta t + (1-\theta)t'$. Next, a similar result holds for the functions h_i. Since they are affine, we may represent $h(x) = Ax + b$, and consequently

$$h(y) = \theta h(x) + (1-\theta)h(x') = \theta v + (1-\theta)v'.$$

We conclude that

$$\theta \begin{pmatrix} u \\ v \\ t \end{pmatrix} + (1-\theta) \begin{pmatrix} u' \\ v' \\ t' \end{pmatrix} = \begin{pmatrix} \theta u + (1-\theta)u' \\ \theta v + (1-\theta)v' \\ \theta t + (1-\theta)t' \end{pmatrix} \in \mathcal{A}$$

with corresponding point $\theta x + (1-\theta)x'$ satisfying the necessary conditions. So, $\mathcal{A}$ is convex.

5. (BV 4.15) In a Boolean linear program, the variable x is constrained to have components equal to zero or one:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \\ & x_i \in \{0,1\}, \quad i = 1, \ldots, n. \end{array}$$

In general, such problems are very difficult to solve, even though the feasible set is finite (containing at most 2^n points). In a general method called relaxation, the constraint that x_i be zero or one is replaced with the linear inequalities $0 \leq x_i \leq 1$:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \\ & 0 \leq x_i \leq 1, \quad i = 1, \ldots, n. \end{array}$$

We refer to this problem as the LP relaxation of the Boolean LP. The LP relaxation is far easier to solve than the original Boolean LP.

(a) Show that the optimal value of the LP relaxation is a lower bound on the optimal value of the Boolean LP. What can you say about the Boolean LP if the LP relaxation is infeasible?

Solution: Let $\mathcal{P}$ be the set of feasible points for the Boolean LP, i.e.,

$$\mathcal{P} = \{x \mid Ax \leq b,\ x \in \{0,1\}^n\}.$$

Likewise, set $\mathcal{P}'$ as the feasible points for the relaxation:

$$\mathcal{P}' = \{x \mid Ax \leq b,\ 0 \leq x \leq 1\}.$$

Clearly $\mathcal{P} \subset \mathcal{P}'$. Therefore minimizing $c^T x$ over $\mathcal{P}'$ gives an objective value at least as small as when minimizing over $\mathcal{P}$. Consequently, solving the relaxation gives a lower bound on the Boolean LP optimal value.

Let p_1 be the optimal objective value for the Boolean LP, and let p_2 be the optimal objective value for the relaxation. If the LP relaxation is infeasible ($p_2 = \infty$), then $p_1 \geq p_2$ and so $p_1 = \infty$ as well. That is, the Boolean LP is infeasible when its relaxation is. Viewed another way, if $\mathcal{P}' = \emptyset$, then since $\mathcal{P} \subset \mathcal{P}'$, $\mathcal{P} = \emptyset$ as well.

(b) It sometimes happens that the LP relaxation has a solution with $x_i \in \{0,1\}$. What can you say in this case?

Solution: In this case, the minimizer x^* for the relaxation is also a minimizer for the Boolean LP. For if not, $p_1 > p_2$ and there would exist a minimizer $y^* \in \mathcal{P}$ with $c^T y^* < c^T x^*$. But y^* is feasible for the relaxation, and this contradicts the optimality of x^*.

```
close all; clc

%%%%%%%%%%%%%% Primal problem and duality bounds %%%%%%%%%%%%%

% primal problem:
% min  x^2 + 1
% s.t. (x-2)(x-4) <= 0


x = linspace(1, 6, 1e3); % the entire feasible region and then some

f0 = @(x) x.^2+1; % vector x, objective function
L = @(x,lam) f0(x) + lam*(x-2).*(x-4); %vector x, scalar lam,
 Lagrangian
xmin = 2;
xmax = 4;

lams = [0.5, 1, 2, 5];
n = length(lams);

% plotting
figure(1)
plot(x,f0(x), 'k-'); hold on
legendstr{1} = 'x^2+1';

for j = 1:n
    lam = lams(j);
    plot(x,L(x,lam), '--')
    legendstr{j+1} = sprintf('L, \\lambda = %g', lam);
end
ylim([0,30])
xlabel x
title('Illustration of primal optimality')

% the feasible region
line([xmin xmin], [0,30], 'color', 'cyan', 'linestyle', ':')
line([xmax xmax], [0,30], 'color', 'cyan', 'linestyle', ':')

% the optimal point
scatter(2,5, 'k', 'filled')
text(2, 9, '(x^*=2, p^*=5)', 'fontsize', 11,'horiz', 'c')

legend(legendstr, 'location','southeast')


%%%%%%%%%%%%%% The dual function %%%%%%%%%%%%%

lams = linspace(-0.9, 10, 1e3);

% dual function for nonnegative dual variable lambda
g = -9*lams.^2./(1+lams) + 8*lams + 1;
```

```
figure(2)
plot(lams, g, 'k-'); hold on
line([-1,-1], [min(g), max(g)], 'linestyle', '--')
xlim([-2, lams(end)])
ylim([min(g), max(g)+5])
xlabel \lambda
ylabel g(\lambda)
title('The Lagrange dual function, g')

%%%%%%%%%%%%%% Perturbed problem -- parameter u %%%%%%%%%%%%%%

% perturbed problem is
% min  x^2 + 1
% s.t. (x-2)(x-4) <= u


f1 = @(x) (x-2).*(x-4);
x = linspace(0,6,1e3);

figure(3)
plot(x,f1(x), 'k-')
xlabel x
ylabel('(x-2)(x-4)')

% optimal primal value
p = @(u) ((3-sqrt(1+u)).^2 + 1) .* (u>=-1 & u<8) ...
    + 1*(u >= 8); % only allow u >= -1

u = linspace(-1,12, 1e3);

figure(4)
plot(u,p(u), 'k-')

xlabel u
ylabel p^*(u)
ylim([0.5,10])
```

Published with MATLAB® R2019b