Note: The purpose of this problem is to explore Lagrange duality, and show how it is useful in creating approximations or proving the quality of approximations to computationally expensive problems. Sometimes, as in the second problem, it will be effectively impossible to solve a problem and so duality provides a welcome (approximate) picture of the problem solution without as much computational cost. Other times, as in the last problem, duality gives a method by which we can show how good our approximate solution is to a hard problem. While the actual regression in the last problem is not classically 'difficult', these polynomial differences in problem difficulty can become seriously important as our data becomes large, which motivates their study in our class.

Boyd 5.28 Prove (without using any linear programming code) that the optimal solution of the LP

$$\text{minimize: } 47x_1 + 93x_2 + 17x_3 - 93x_4$$

$$\text{subj. to: } \begin{bmatrix} -1 & -6 & 1 & 3 \\ -1 & -2 & 7 & 1 \\ 0 & 3 & -10 & -1 \\ -6 & -11 & -2 & 12 \\ 1 & 6 & -1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \preceq \begin{bmatrix} -3 \\ 5 \\ -8 \\ -7 \\ 4 \end{bmatrix}$$

is unique, and given by $x^\star = (1, 1, 1, 1)$.

■

Boyd 5.13 A Boolean linear program is an optimization problem of the form

$$\begin{aligned} \text{minimize: } & c^* x \\ \text{subj. to: } & Ax \preceq b, \\ & x_i \in \{0, 1\}. \end{aligned}$$

Boolean linear programming is NP-complete and, as such, difficult to solve. As an example application, we might wish to perform an ℓ_1 penalty regression that used no more than 10 features of say 500 available features so as to increase prediction speed; this can be formulated as a boolean LP. Despite the algorithmic challenge, we will derive a formulation which will give an efficient lower bound on the optimal solution above.

We can rewrite the above LP as

$$\begin{aligned} \text{minimize: } & c^* x \\ \text{subj. to: } & Ax \preceq b, \\ & x_i(1 - x_i) = 0. \end{aligned}$$

Find the Lagrange dual of this problem and show that it is convex. Thus, the optimal dual value can be efficiently solved and gives a lower bound for the boolean LP problem.

■

Boyd 5.6 Consider the Chebyshev or ℓ_∞-norm regression problem

$$\text{minimize } \|Ax - b\|_\infty,$$

where $A \in \mathbb{R}^{m \times n}$ and rank $A = n$. Let x_{ch} denote an optimal solution (there may be many). Solving the Chebyshev regression problem is slower in general (via linear programming) than solving the least squares problem $x_{\mathrm{ls}} = \arg\min \|Ax - b\|_2$. We will check how good of an approximation it is to return x_{ls} instead of an actual solution x_{ch} for the Chebyshev regression problem. Put another way, we will find bounds on how much larger $\|Ax_{\mathrm{ls}} - b\|_\infty$ is than $\|Ax_{\mathrm{ch}} - b\|_\infty$.

(a) Prove the lower bound

$$\|Ax_{\mathrm{ls}} - b\|_\infty \le \sqrt{m}\|Ax_{\mathrm{ch}} - b\|_\infty$$

using standard equivalence-of-norm inequalities.

(b) The dual of the ℓ_∞ regression problem is

$$\text{maximize: } b^*\nu$$
$$\text{subj. to: } \|\nu\|_1 \le 1$$
$$A^*\nu = 0.$$

Thus ν in the kernel of A^* with ℓ_1 norm at most one gives a lower bound $b^*\nu$ on $\|Ax_{\mathrm{ch}} - b\|_\infty$. Denote the least-squares residual by $r_{\mathrm{ls}} = b - Ax_{\mathrm{ls}}$. Assuming $r_{\mathrm{ls}} \ne 0$, show that

$$\hat{\nu} = -r_{\mathrm{ls}}/\|r_{\mathrm{ls}}\|_1, \text{ and } \widetilde{\nu} = r_{\mathrm{ls}}/\|r_{\mathrm{ls}}\|_1$$

are both feasible in the dual problem. Which corresponding lower bound on $\|Ax_{\mathrm{ch}} - b\|_\infty$ is better? How do these compare with the bound derived in part (a)?

■