The University of Texas at Austin

Optimization
Homework 4

Patrick Brown
EID: pbb468

Instructors: Constantine Caramanis, Sujay Sanghavi

Problem 1: Unit eigenvectors of stochastic matrices

One of the most powerful modeling ideas in stochastics and probability, used in Machine Learning, Reinforcement Learning, and many many other fields of mathematics, engineering and applied science, is the concept of the Markov Chain: this is a stochastic process whose state makes past and future independent. In other words, if you want to know what will happen tomorrow (and what you should do today), you only need to look at where you are, and not how you got there.

A foundational fact that is extremely useful in the study of Markov Chains, is the fact that every finite state Markov chain has an invariant probability distribution. It is beyond the scope of this class (and certainly this homework!) to discuss this any further; however, as you will see, we have already developed enough tools to actually prove this, using optimization and duality.

We say that an $n \times n$ matrix P, with entries p_{ij}, is stochastic if all of its entries are nonnegative and

$$\sum_{j=1}^{n} p_{ij} = 1, \quad \forall i,$$

that is, the sum of the entries of each row is equal to 1.

Use duality to show that if P is a stochastic matrix, then the system of equations,

$$P^T \lambda = \lambda, \quad \lambda \geq 0,$$

has a nonzero solution. (Note that the vector λ can be normalized so that its components sum to one - this normalized vector is the invariant probability distribution.)

1 Solution to Problem 1

To show that the system of equations $P^T \lambda = \lambda$, with $\lambda \geq 0$, has a nonzero solution when P is a stochastic matrix, we can use linear programming duality. Here's how:

1.1 Step 1: Formulate the Primal Linear Program

Consider the following primal linear program (LP):

$$\begin{aligned} \text{Maximize} \quad & e^T \lambda \\ \text{Subject to} \quad & (P^T - I)\lambda = 0 \\ & \lambda \geq 0 \end{aligned}$$

where e is the vector of ones, and I is the identity matrix. The objective is to maximize the sum of the components of λ, under the constraint $P^T \lambda = \lambda$ and $\lambda \geq 0$.

1.2 Step 2: Formulate the Dual Linear Program

The dual of the above primal LP is:

$$
\begin{aligned}
\text{Minimize} \quad & 0 \\
\text{Subject to} \quad & (P - I)y + s \geq e \\
& s \geq 0
\end{aligned}
$$

Here, y and s are the dual variables corresponding to the equality and inequality constraints of the primal LP, respectively.

1.3 Step 3: Analyze the Dual Feasibility

Notice that the dual objective function is zero, and any feasible solution would yield an objective value of zero. However, let's investigate the feasibility of the dual constraints.

Assume that there exists y and $s \geq 0$ satisfying $(P - I)y + s \geq e$. Since P is stochastic, $(P - I)$ has non-positive diagonal entries and non-negative off-diagonal entries. Adding $s \geq 0$ cannot make the left-hand side $(P - I)y + s$ greater than or equal to e unless y and s are unreasonably large or s is infinite, which is not possible.

Therefore, the dual LP is **infeasible**.

1.4 Step 4: Conclude Using Duality

By the **Strong Duality Theorem** in linear programming, if the dual LP is infeasible, then the primal LP is either unbounded or infeasible. Since $\lambda = 0$ is a feasible solution to the primal LP (though it gives an objective value of zero), the primal LP is **feasible**.

Given that the primal LP is feasible and the dual LP is infeasible, the primal LP must be **unbounded**. This means that we can find $\lambda \geq 0$ satisfying $(P^T - I)\lambda = 0$ such that $e^T \lambda$ can be made arbitrarily large.

1.5 Step 5: Conclusion

An unbounded primal problem implies that there exists a nonzero $\lambda \geq 0$ satisfying $(P^T - I)\lambda = 0$, i.e., $P^T \lambda = \lambda$.

Thus, using duality, we've shown that the system $P^T \lambda = \lambda$, $\lambda \geq 0$ has a nonzero solution. This vector λ can be normalized so that its components sum to one, making it the invariant probability distribution of the Markov chain defined by P.

Final Answer:

Yes; using LP duality, we show that since the dual of an appropriate LP is infeasible, the system $P^T \lambda = \lambda$, $\lambda \geq 0$ has a nonzero solution.

Problem 2: Duality in piecewise linear convex optimization

Consider the problem of minimizing $\max_{i=1,\ldots,m}(a_i^T x - b_i)$ over all $x \in \mathbb{R}^n$. Let v be the value of the optimal cost, assumed finite. Let A be the matrix with rows $a_1, \ldots, a_m$, and let b be the vector with components $b_1, \ldots, b_m$.

(a) Consider any vector $p \in \mathbb{R}^m$ that satisfies $A^T p = 0$, $p \geq 0$, and $\sum_{i=1}^m p_i = 1$. Show that $-p^T b \leq v$.

(b) In order to obtain the best possible lower bound of the form considered in part (a), we form the linear programming problem

$$
\begin{aligned}
\text{maximize} \quad & -p^T b \\
\text{subject to} \quad & A^T p = 0 \\
& p^T e = 1 \\
& p \geq 0,
\end{aligned}
$$

where e is the vector with all components equal to 1. Show that the optimal cost in this problem is equal to v.

2 Solution to Problem 2

2.1 Part (a)

To prove that $-p^T b \leq v$, we will show that for any $x \in \mathbb{R}^n$:

$$\max_{i=1,\ldots,m} (a_i^T x - b_i) \geq -p^T b.$$

Since $v = \min_{x \in \mathbb{R}^n} \max_i (a_i^T x - b_i)$, it will follow that $-p^T b \leq v$.

Proof:

1. **Start with the definition of the maximum:**

   For any $x \in \mathbb{R}^n$:

   $$\max_{i=1,\ldots,m} (a_i^T x - b_i) \geq a_i^T x - b_i \quad \text{for all } i.$$

2. **Take a convex combination using p:**

   Multiply both sides by $p_i \geq 0$ and sum over i:

   $$\sum_{i=1}^{m} p_i \max_{j=1,\ldots,m} (a_j^T x - b_j) \geq \sum_{i=1}^{m} p_i (a_i^T x - b_i)$$

3. **Use the properties of p:**

   Since $\sum_{i=1}^{m} p_i = 1$, $A^T p = 0$, and $p \geq 0$:

   $$\max_{i=1,\ldots,m} (a_i^T x - b_i) \geq \sum_{i=1}^{m} p_i a_i^T x - \sum_{i=1}^{m} p_i b_i = (A^T p)^T x - p^T b = -p^T b$$

4. **Conclude:**

   We have shown that for any $x \in \mathbb{R}^n$:

   $$\max_{i=1,\ldots,m} (a_i^T x - b_i) \geq -p^T b$$

   Therefore, $-p^T b \leq v$.

2.2 Part (b)

Primal Problem:

$$\begin{aligned} \text{minimize} \quad & t \\ \text{subject to} \quad & a_i^T x - b_i - t \leq 0, \quad i = 1, \ldots, m. \end{aligned}$$

Dual Problem:

- Dual variables: $p \geq 0$.

- Dual objective: Maximize $(-b)^T p$.

- Dual constraints:

$$\begin{cases} A^T p = 0, \\ p^T e = 1. \end{cases}$$

The dual problem simplifies to:

$$
\begin{aligned}
\text{maximize} \quad & -p^T b \\
\text{subject to} \quad & A^T p = 0 \\
& p^T e = 1 \\
& p \geq 0.
\end{aligned}
$$

Invoke strong duality:

- The primal problem is a linear program and satisfies the conditions for strong duality (it is feasible and bounded).

- Therefore, the optimal values of the primal and dual problems are equal:

$$
v = \min_x \max_i (a_i^T x - b_i) = \max_{p \geq 0, A^T p = 0, p^T e = 1} -p^T b.
$$

Conclusion:

The optimal cost of the given linear programming problem is equal to v, the minimal value of the original optimization problem.

Problem 3: Optimality in Linear Programming with Symmetric Constraints

Let A be a symmetric square matrix. Consider the linear programming problem

$$
\begin{aligned}
\text{minimize} \quad & c^T x \\
\text{subject to} \quad & Ax \geq c \\
& x \geq 0.
\end{aligned}
$$

Prove that if x^* satisfies $Ax^* = c$ and $x^* \geq 0$, then x^* is an optimal solution.

3 Solution to Problem 3

To prove that x^* is an optimal solution, we will use the duality theory of linear programming.

3.1 Primal Problem:

$$
\begin{aligned}
\text{Minimize} \quad & c^T x \\
\text{Subject to} \quad & Ax \geq c \\
& x \geq 0.
\end{aligned}
$$

3.2 Dual Problem:

The dual of the given primal problem is:

$$
\begin{aligned}
\text{Maximize} \quad & c^T y \\
\text{Subject to} \quad & A^T y \leq c \\
& y \geq 0.
\end{aligned}
$$

Note: Since A is symmetric, $A^T = A$.

3.3 Constructing Dual Solution:

Let $y^* = x^*$.

Since $x^* \geq 0$, it follows that $y^* \geq 0$.

Check the dual constraints:

$$Ay^* = Ax^* = c \leq c,$$

which implies $Ay^* \leq c$. Thus, y^* is feasible for the dual problem.

3.4 Objective Values:

Calculate the objective values at x^* and y^*:

- Primal objective value: $c^T x^*$.

- Dual objective value: $c^T y^* = c^T x^*$.

3.5 Conclusion:

Since the objective values are equal and both x^* and y^* are feasible for their respective problems, by the **Strong Duality Theorem**, x^* is an optimal solution to the primal problem.

Therefore, x^* is an optimal solution.

Problem 4: Robust Linear Programming

Consider the following problem:

$$
\begin{aligned}
\min: &\quad c_1 x_1 + c_2 x_2 + c_3 x_3 \\
\text{s.t.}: &\quad Ax \geq b, \quad \forall b \in \mathcal{U},
\end{aligned}
$$

where the uncertainty set $\mathcal{U}$ is the unit box in m dimensions, i.e.,

$$\mathcal{U} = \{b : 0 \leq b_i \leq 1, i = 1, ..., m, \sum_{i=1}^{m} b_i \leq 1\}.$$

Show that this is equivalent to the LP

$$
\begin{aligned}
\min: &\quad c_1 x_1 + c_2 x_2 + c_3 x_3 \\
\text{s.t.}: &\quad Ax \geq \hat{b},
\end{aligned}
$$

where

$$\hat{b} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

4 Solution to Problem 4

The goal is to show that the robust linear programming problem:

$$\min \quad c_1 x_1 + c_2 x_2 + c_3 x_3$$
$$\text{s.t.} \quad Ax \geq b, \quad \forall b \in \mathcal{U},$$

where the uncertainty set $\mathcal{U}$ is defined as:

$$\mathcal{U} = \left\{ b \in \mathbb{R}^m : 0 \leq b_i \leq 1 \text{ for } i = 1, \ldots, m, \quad \sum_{i=1}^{m} b_i \leq 1 \right\},$$

is equivalent to the linear program:

$$\min \quad c_1 x_1 + c_2 x_2 + c_3 x_3$$
$$\text{s.t.} \quad Ax \geq \hat{b},$$

4.1 Explanation:

4.1.1 1. Understanding the Uncertainty Set $\mathcal{U}$:

The set $\mathcal{U}$ is defined by the constraints:

- Each component b_i satisfies $0 \leq b_i \leq 1$.

- The sum of all components satisfies $\sum_{i=1}^{m} b_i \leq 1$.

- Since $\sum_{i=1}^{m} b_i \leq 1$, only one b_i can be 1 at a time, with the others being zero.

However, to ensure the constraints hold **simultaneously** for all possible $b \in \mathcal{U}$, we must consider all such worst-case scenarios.

4.1.2 2. Deriving the Equivalent Constraints:

To guarantee that $Ax \geq b$ for all $b \in \mathcal{U}$, it is sufficient to impose:

$$A_i x \geq 1, \quad \text{for all } i = 1, \ldots, m.$$

This is because:

- For any i, the maximum b_i can be is 1.

- If $A_i x \geq 1$ holds for all i, then $A_i x \geq b_i$ will hold for any $b_i \leq 1$.

4.1.3 3. Conclusion:

By setting $\hat{b} = [1, 1, \ldots, 1]^\top$, the robust constraints simplify to:

$$Ax \geq \hat{b}.$$

Therefore, the original robust linear programming problem is equivalent to the standard linear program with the right-hand side vector $\hat{b}$:

$$\min \quad c^\top x$$
$$\text{s.t.} \quad Ax \geq \hat{b}.$$

4.2   Summary:

The equivalence holds because the worst-case scenario for the constraints under the given uncertainty set $\mathcal{U}$ occurs when each b_i is at its maximum value of 1. Ensuring that $Ax \geq \hat{b}$ accounts for all possible b in $\mathcal{U}$, thus making the robust problem equivalent to the standard linear program with $Ax \geq \hat{b}$.