ENGG 5501: Foundations of Optimization	**2016–17 First Term**

Handout 4: Some Applications of Linear Programming

Instructor: Anthony Man–Cho So — October 12, 2016

1 Introduction

The theory of LP has found many applications in various disciplines. In this lecture, we will consider some of those applications and see how the machineries developed in previous lectures can be used to obtain some interesting results.

2 Arbitrage–Free Asset Pricing

Consider a market in which n different assets are traded. We are interested in the performance of the assets in one period, which naturally is influenced by the events during that period. For simplicity's sake, let us assume that there are m possible states at the end of the period. Now, suppose that for every unit of asset $i \in \{1, \ldots, n\}$ owned, we can receive a payoff of r_{si} dollars if state $s \in \{1, \ldots, m\}$ is realized at the end of the period. In other words, the payoff information can be captured by the following $m \times n$ payoff matrix R:

$$R = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mn} \end{bmatrix}.$$

Let $x_i \in \mathbb{R}$ be number of units of asset i held. Note that if $x_i \geq 0$, then we can get $r_{si}x_i$ dollars at the end of the period if state s is realized. On the other hand, if $x_i < 0$, then we are in a "short" position with respect to asset i. This means that we are selling $|x_i|$ units of asset i at the beginning of the period, with a promise to buy them back at the end of the period. In particular, we need to pay $r_{si}|x_i|$ dollars if state s is realized, which of course is equivalent to receiving a payoff of $r_{si}x_i$ dollars.

Clearly, given an initial portfolio $x = (x_1, \ldots, x_n)$, our wealth at the end of the period is

$$w_s = \sum_{i=1}^{n} r_{si}x_i$$

if state s is realized. Upon letting $w = (w_1, \ldots, w_m)$, we see that $w = Rx$. Now, one of the most fundamental problems in finance is to determine the prices of the assets at the beginning of the period. To formalize the problem, let p_i be the price of asset i at the beginning of the period, and set $p = (p_1, \ldots, p_n)$. Then, the cost of acquiring the portfolio x is given by $p^T x$. A fundamental assumption in finance theory is that the prices should not give rise to **arbitrage opportunities**; i.e., no investor should be able to get a non–negative payoff out of a negative investment. In other words, any portfolio that guarantees a non–negative payoff in every state must be valuable to investors and hence must have a non–negative price. The arbitrage–free assumption can be expressed as follows:

$$\text{if } Rx \geq \mathbf{0} \text{ then } p^T x \geq 0. \tag{1}$$

Now, given the payoff matrix R, do there exist arbitrage–free prices of the assets? As the following result shows, such a question can be answered by Farkas' lemma:

Theorem 1 *There is no arbitrage opportunity (i.e., condition (1) holds) if and only if there exists a $q \in \mathbb{R}^m_+$ such that the price of asset $i \in \{1, \ldots, n\}$ is given by*

$$p_i = \sum_{s=1}^{m} q_s r_{si}.$$

Proof The arbitrage–free condition (1) is equivalent to the statement that the following system has no solution:

$$Rx \geq \mathbf{0}, \quad p^T x < 0. \tag{2}$$

By Farkas' lemma, there exists a $q \in \mathbb{R}^m_+$ such that $R^T q = p$, which is precisely the condition claimed in the theorem's statement. □

We should point out that in the above market model, arbitrage opportunities can be detected by simply solving a linear programming problem. Indeed, it suffices to determine the feasibility of the following linear system:

$$Rx \geq \mathbf{0}, \quad p^T x = -1. \tag{3}$$

This follows from the fact that we can always scale a solution to (2) to get a solution to (3).

For further reading on the application of optimization theory to economics and finance, we refer the reader to [3, 2].

3 An Application in Cooperative Game Theory

Consider a set $\mathcal{N} = \{1, \ldots, n\}$ of n players. Let $v : 2^{\mathcal{N}} \to \mathbb{R}_+$ be a **value function** (also known as a **characteristic function**); i.e., for $S \subset \mathcal{N}$, the value $v(S)$ is the **worth** of the **coalition** S. In other words, $v(S)$ represents the total worth that the members of S can earn without any help from the players outside S. By default, we set $v(\emptyset) = 0$. Naturally, the total worth $v(\mathcal{N})$ is shared among all the players, and we would like to know how to share it so that no coalition has the incentive to deviate and obtain an outcome that is better for all of its members. Specifically, consider an **allocation vector** $x \in \mathbb{R}^n_+$, where $x_i \geq 0$ represents the payoff to player i, where $i = 1, \ldots, n$. We say that a coalition S can improve upon an allocation x iff $v(S) > x(S) \equiv \sum_{i \in S} x_i$. Equivalently, S can improve upon x iff there exists some allocation $y \in \mathbb{R}^n_+$ such that $y(S) = v(S)$ and $y_i > x_i$ for all $i \in S$. Finally, we say that an allocation x is in the **core** if

$$x(\mathcal{N}) = v(\mathcal{N}), \quad x(S) \geq v(S) \qquad \text{for all } S \subset \mathcal{N}.$$

The first condition simply indicates that we are dividing the total worth $v(\mathcal{N})$ among the players. Note that if an allocation x such that $x(\mathcal{N}) = v(\mathcal{N})$ is not in the core, then there is some coalition S in which all of its members can do strictly better than in x by cooperating together and dividing the worth $v(S)$ among themselves. On the other hand, if x is in the core, then the allocation is **stable**, and no coalition has the incentive to deviate.

Now, a natural question is whether the core always exists. Before we address that question, let us consider an example.

Example 1 *Suppose that an expedition of n people has discovered treasure in the mountains. However, it takes two people to carry out each piece of treasure, in which case the value of the treasure is equally shared between the two. To model this game, we define the value function via $v(S) = \lfloor |S|/2 \rfloor$ for any $S \subset \mathcal{N}$. Consider first the case where n is even. Then, the allocation $(1/2, 1/2, \ldots, 1/2)$ is* stable*. Now, suppose that $n = 3$. Note that the allocation $(1/2, 1/2, 0)$ does not belong to the core, since if the third player pairs up with, say, the second player and promises to give her $1/2 + \epsilon$ for some $\epsilon > 0$ while keeping $1/2 - \epsilon$ for herself, then both of them are better off. Similarly, the allocation $(1/3, 1/3, 1/3)$ is not in the core, since two players can each get $1/2$ if they form a coalition. Indeed, it is not hard to show that when $n \geq 3$ is odd, then the core is empty.*

From the previous example, we see that given an arbitrary coalition game $(\mathcal{N}, v)$, the core may be empty. It turns out that the problem of characterizing those coalition games that have non–empty cores can be resolved using LP.

Theorem 2 (Bondareva–Shapley; see [6]) *The coalition game $(\mathcal{N}, v)$ has a non–empty core iff the following* **balancedness condition** *is satisfied: For any set of weights $\{y_S\}_S$ such that $y_S \geq 0$ and $\sum_{S:i\in S} y_S = 1$ for $i = 1, \ldots, n$, we have $\sum_S y_S v(S) \leq v(\mathcal{N})$.*

Proof Consider the following LP:

$$\begin{array}{lll} \text{minimize} & x(\mathcal{N}) & \\ \text{subject to} & x(S) \geq v(S) & \text{for all } S \subset \mathcal{N}. \end{array} \tag{4}$$

The dual of Problem (4) is given by

$$\begin{array}{lll} \text{maximize} & \displaystyle\sum_S y_S v(S) & \\ \text{subject to} & \displaystyle\sum_{S:i\in S} y_S = 1 & \text{for } i = 1, \ldots, n, \\ & y \geq \mathbf{0}. & \end{array} \tag{5}$$

Since Problem (4) is feasible and its optimal value is lower bounded by $v(\mathcal{N})$, by Corollary 1 of Handout 3 and the LP strong duality theorem, there exist optimal solutions x^* and y^* to (4) and (5), respectively, with

$$x^*(\mathcal{N}) = \sum_S y_S^* v(S).$$

Now, observe that the allocation x^* is in the core iff $x^*(\mathcal{N}) \leq v(\mathcal{N})$. The latter is equivalent to the fact that the balancedness condition is satisfied, and the proof is completed. □

For further reading on game theory, we refer the reader to [5]. Lloyd S. Shapley was awarded The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2012 "for the theory of stable allocations and the practice of market design"; see `http://www.nobelprize.org/nobel_prizes/economics/laureates/2012/` for details.

4 An Approximation Algorithm for Vertex Cover

The theory of LP has been employed very successfully in the design of approximation algorithms in recent years (see, e.g., [7]). We say that an algorithm $\mathcal{A}$ is an α–**approximation algorithm** for a minimization problem $\mathcal{P}$ if for every instance I of $\mathcal{P}$, it delivers in polynomial time a solution whose objective value is at most α times the optimal value. Clearly, we have $\alpha \geq 1$, and the closer it is to 1, the better. In a similar fashion, one may define the notion of an α–approximation algorithm for a maximization problem. In this section we study the so–called *vertex cover* problem and see how the theory of LP can be used to obtain a 2–approximation algorithm for it.

To begin, consider a simple undirected graph $G = (V, E)$, where each vertex $v_i \in V$ has an associated cost $c_i \in \mathbb{R}_+$. A **vertex cover** of G is a subset $S \subset V$ such that for every edge $(v_i, v_j) \in E$, at least one of the endpoints belongs to S. We are interested in finding a vertex cover S of G of minimal cost.

Now, let $x_i \in \{0, 1\}$ be a binary variable indicating whether v_i belongs to the vertex cover S or not (i.e., $x_i = 1$ iff $v_i \in S$). Then, the minimum–cost vertex cover problem can be formulated as the following integer program:

$$\begin{array}{rcll} v^* & = & \text{minimize} & c^T x = \sum_{i=1}^{|V|} c_i x_i \\ & & \text{subject to} & x_i + x_j \geq 1 \qquad \text{for } (v_i, v_j) \in E, \\ & & & x \in \{0,1\}^{|V|}. \end{array} \tag{6}$$

Due to the presence of the integer constraint $x \in \{0, 1\}^{|V|}$, Problem (6) is generally difficult to solve. One intuitive strategy for circumventing this difficulty is to replace the integer constraint $x \in \{0, 1\}^{|V|}$ with the linear constraint $\mathbf{0} \leq x \leq e$. Using the fact that $c \geq \mathbf{0}$, it is not hard to show that the resulting problem is equivalent to the following LP, which is called an *LP relaxation* of Problem (6):

$$\begin{array}{rcll} v_r^* & = & \text{minimize} & c^T x \\ & & \text{subject to} & x_i + x_j \geq 1 \quad \text{for } (v_i, v_j) \in E, \\ & & & x \geq \mathbf{0}. \end{array} \tag{7}$$

Clearly, we have $v_r^* \leq v^*$. Suppose that x' is an optimal solution to Problem (7). It is then natural to ask whether we can convert x' into a solution x'' that is feasible for Problem (6) and satisfies $c^T x'' \leq \alpha v_r^*$ for some $\alpha > 0$. Note that if this is possible, then we would obtain an α–approximation algorithm for the minimum–cost vertex cover problem because $v_r^* \leq v^*$. As it turns out, the answer to the above question is indeed affirmative. The key to proving this is the following theorem:

Theorem 3 *(cf. [4]) Let $P \subset \mathbb{R}^{|V|}$ be the polyhedron defined by the following system:*

$$\left\{ \begin{array}{rcll} x_i + x_j & \geq & 1 & \text{for } (v_i, v_j) \in E, \\ x & \geq & \mathbf{0}. \end{array} \right.$$

Suppose that x is an extreme point of P. Then, we have $x_i \in \{0, 1/2, 1\}$ for $i = 1, \ldots, |V|$.

Proof Let $x \in P$ and consider the sets

$$\begin{array}{rcl} U_{-1} & = & \{ i \in \{1, \ldots, |V|\} : x_i \in (0, 1/2) \}, \\ U_1 & = & \{ i \in \{1, \ldots, |V|\} : x_i \in (1/2, 1) \}. \end{array}$$

For $i = 1, \ldots, |V|$ and $k \in \{-1, 1\}$, define

$$y_i = \begin{cases} x_i + k\epsilon & \text{if } i \in U_k, \\ x_i & \text{otherwise} \end{cases}, \quad z_i = \begin{cases} x_i - k\epsilon & \text{if } i \in U_k, \\ x_i & \text{otherwise.} \end{cases}$$

By definition, we have $x = (y + z)/2$. If either U_{-1} or U_1 is non–empty, then we may choose $\epsilon > 0$ to be sufficiently small so that $y, z \in P$, and that x, y, z are all distinct. It follows that $U_k = \emptyset$ for $k \in \{-1, 1\}$ if x is an extreme point of P. □

Corollary 1 *There exists a* 2*–approximation algorithm for the minimum–cost vertex cover problem.*

Proof We first solve the LP (7) and obtain an optimal extreme point solution x'. Now, by Theorem 3, all the entries of x' belong to $\{0, 1/2, 1\}$. Hence, the vector x'' defined by

$$x''_i = \begin{cases} x'_i & \text{if } x'_i = 0 \text{ or } 1, \\ 1 & \text{if } x'_i = 1/2 \end{cases} \qquad \text{for } i = 1, \ldots, |V|$$

is feasible for Problem (6). Moreover, it has objective value $c^T x'' \le 2c^T x' = 2v_r^* \le 2v^*$. This completes the proof. □

5 Blind Separation of Non–Negative Sources

In various image processing applications, a problem of fundamental interest is that of separating non–negative source signals in a blind fashion. For simplicity, consider the following linear mixture model:

$$x[\ell] = As[\ell] \quad \text{for } \ell = 1, \ldots, L, \tag{8}$$

where $s[\ell] \in \mathbb{R}^n_+$ is the ℓ–th source vector, $x[\ell] \in \mathbb{R}^m$ is the ℓ–th observation vector, and $A \in \mathbb{R}^{m\times n}$ is a mixing matrix describing the input–output relationship. Our goal here is to extract the source vectors $s[1], \ldots, s[L] \in \mathbb{R}^n$ from the observation vectors $x[1], \ldots, x[\ell] \in \mathbb{R}^m$ without knowledge of the mixing matrix $A \in \mathbb{R}^{m\times n}$. Note that such a task is not well–defined. For instance, if the pair $\left(\{s[\ell]\}_{\ell=1}^L, A\right)$ satisfies (8), then so does the pair $\left(\{s[\ell]/c\}_{\ell=1}^L, cA\right)$ for any constant $c > 0$. Thus, it is necessary to impose additional assumptions on the model (8). Towards that end, let us first rewrite (8) as

$$x^i = \sum_{j=1}^n a_{ij} s^j \quad \text{for } i = 1, \ldots, m, \tag{9}$$

where $x^i = (x^i[1], \ldots, x^i[L]) \in \mathbb{R}^L$ is the i–th observed signal and $s^j = (s^j[1], \ldots, s^j[L]) \in \mathbb{R}^L_+$ is the signal from the j–th source. We shall make the following assumptions:

(a) The mixing matrix $A \in \mathbb{R}^{m\times n}$ has full column rank (in particular, $m \ge n$) and satisfies

$$\sum_{j=1}^n a_{ij} = 1 \quad \text{for } i = 1, \ldots, m.$$

Moreover, the number of observations L satisfies $L \gg m$.

(b) Each source signal is *locally dominant*; i.e., for each source $j \in \{1, \dots, n\}$, there exists an (unknown) undex $\ell(j) \in \{1, \dots, L\}$ such that $s^j[\ell(j)] > 0$ and $s^k[\ell(j)] = 0$ for all $k \neq j$.

The above assumptions are satisfied in a wide variety of settings; see, e.g., [1] for a more detailed discussion. Now, observe that from Assumption (a), we have $x^i \in \mathrm{aff}(\{s^1, \dots, s^n\})$ for $i = 1, \dots, m$. In fact, more can be said:

Proposition 1 *Under Assumption (a), we have* $\mathrm{aff}(\{x^1, \dots, x^m\}) = \mathrm{aff}(\{s^1, \dots, s^n\})$.

Proof Suppose that $x \in \mathrm{aff}(\{x^1, \dots, x^m\})$. Then, there exist $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ such that

$$x = \sum_{i=1}^{m} \alpha_i x^i \quad \text{and} \quad \sum_{i=1}^{m} \alpha_i = 1.$$

Substituting this into (9) yields

$$x = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i a_{ij} s^j = \sum_{j=1}^{n} \beta_j s^j,$$

where $\beta_j = \sum_{i=1}^{m} \alpha_i a_{ij}$, for $j = 1, \dots, n$. Using Assumption (a), we have

$$\sum_{j=1}^{n} \beta_j = \sum_{i=1}^{m} \alpha_i \left(\sum_{j=1}^{n} a_{ij} \right) = 1.$$

It follows that $x \in \mathrm{aff}(\{s^1, \dots, s^n\})$.

Conversely, suppose that $x \in \mathrm{aff}(\{s^1, \dots, s^n\})$. Then, there exist $\beta_1, \dots, \beta_n \in \mathbb{R}$ such that

$$x = \sum_{j=1}^{n} \beta_j s^j \quad \text{and} \quad \sum_{j=1}^{n} \beta_j = 1.$$

Consider now the following system of linear equations in $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m$:

$$\beta_j = \sum_{i=1}^{m} \alpha_i a_{ij} \quad \text{for } j = 1, \dots, n. \tag{10}$$

By Assumption (a), we have $m \geq n$, which implies that the above system is solvable. Upon summing (10) over $j = 1, \dots, n$, we have

$$1 = \sum_{j=1}^{n} \beta_j = \sum_{i=1}^{m} \alpha_i \left(\sum_{j=1}^{n} a_{ij} \right) = \sum_{i=1}^{m} \alpha_i.$$

This shows that $x \in \mathrm{aff}(\{x^1, \dots, x^m\})$, as desired. □

The upshot of Proposition 1 is that although we do not know the source vectors $s^1, \dots, s^n \in \mathbb{R}^L_+$, their affine hull is completely determined by the observed vectors $x^1, \dots, x^m \in \mathbb{R}^L$. Such an observation can help us in recovering the source vectors $s^1, \dots, s^n \in \mathbb{R}^L_+$. Indeed, consider the polyhedron

$$\mathcal{P} = \mathrm{aff}(\{s^1, \dots, s^n\}) \cap \mathbb{R}^L_+. \tag{11}$$

Since $s^i \in \mathbb{R}^L_+$, we have $s^i \in \mathcal{P}$ for $i = 1, \dots, n$. As the following result shows, the source vectors can be recovered by considering the vertices of $\mathcal{P}$:

Proposition 2 *Under Assumption (b), we have $\mathcal{P} = \text{conv}(\{s^1, \dots, s^n\})$. Moreover, the vertices of $\mathcal{P}$ are $\{s^1, \dots, s^n\}$.*

Proof Suppose that $x \in \mathcal{P}$. Then, there exist $\beta_1, \dots, \beta_n \in \mathbb{R}$ such that

$$\mathbf{0} \le x = \sum_{j=1}^n \beta_j s^j \quad \text{and} \quad \sum_{j=1}^n \beta_j = 1.$$

For each $j \in \{1, \dots, n\}$, we have $0 \le x[\ell(j)] = \beta_j s^j[\ell(j)]$ by Assumption (b), which implies that $\beta_j \ge 0$. Thus, we have $x \in \text{conv}(\{s^1, \dots, s^n\})$.

Conversely, suppose that $x \in \text{conv}(\{s^1, \dots, s^n\})$. Since $s^1, \dots, s^n \in \mathbb{R}^L_+$, it is clear that $x \in \mathcal{P}$.

In view of the above, the vertices of $\mathcal{P}$ must belong to the set $\{s^1, \dots, s^n\}$. Thus, to complete the proof of Proposition 2, it remains to show that $s^1, \dots, s^n$ are all vertices of $\mathcal{P}$. Towards that end, let us fix $k \in \{1, \dots, n\}$ and suppose that $s^k = \theta x^1 + (1-\theta)x^2$, where $x^1, x^2 \in \mathcal{P}$ and $\theta \in (0,1)$. Then, there exist $\alpha^1_1, \dots, \alpha^1_n \in \mathbb{R}_+$ and $\alpha^2_1, \dots, \alpha^2_n \in \mathbb{R}_+$ such that

$$s^k = \sum_{j=1}^n \left(\theta\alpha^1_j + (1-\theta)\alpha^2_j \right) s^j \quad \text{and} \quad \sum_{j=1}^n \alpha^1_j = \sum_{j=1}^n \alpha^2_j = 1.$$

Now, by Assumption (b), we have

$$s^k[\ell(k)] = \left(\theta\alpha^1_k + (1-\theta)\alpha^2_k \right) s^k[\ell(k)],$$

which implies that $\theta\alpha^1_k + (1-\theta)\alpha^2_k = 1$. This is possible if and only if $\alpha^1_k = \alpha^2_k = 1$, or equivalently, $x^1 = x^2 = s^k$. It follows that s^k is a vertex of $\mathcal{P}$, as desired. □

To obtain a representation of $\mathcal{P}$ that is more amenable to computation, we first observe that $\dim\left(\text{aff}(\{s^1, \dots, s^n\})\right) = n-1$. Thus, by Proposition 1 and Assumption (b), there are $n-1$ linearly independent vectors in the collection $\{x^i - x^1\}_{i=2}^m$. Now, let $v^1, \dots, v^{L-n+1} \in \mathbb{R}^L$ be a basis of $\text{span}(\{x^2 - x^1, \dots, x^m - x^1\})^\perp$, which can be computed by the Gram–Schmidt orthogonalization procedure. Then, we have

$$\text{aff}(\{x^1, \dots, x^m\}) = \left\{ x \in \mathbb{R}^L : (v^i)^T x = (v^i)^T x^1 \quad \text{for } i = 1, \dots, L-n+1 \right\},$$

which, together with (11), implies that

$$\mathcal{P} = \left\{ x \in \mathbb{R}^L_+ : (v^i)^T x = (v^i)^T x^1 \quad \text{for } i = 1, \dots, L-n+1 \right\}.$$

References

[1] T.-H. Chan, W.-K. Ma, C.-Y. Chi, and Y. Wang. A Convex Analysis Framework for Blind Separation of Non–Negative Sources. *IEEE Transactions on Signal Processing*, 56(10):5120–5134, 2008.

[2] G. Cornuejols and R. Tütüncü. *Optimization Methods in Finance*, volume 5 of *Mathematics, Finance and Risk.* Cambridge University Press, Cambridge, 2007.

[3] A. K. Dixit. *Optimization in Economic Theory.* Oxford University Press, New York, second edition, 1990.

[4] G. L. Nemhauser and L. E. Trotter, Jr. Properties of Vertex Packing and Independence System Polyhedra. *Mathematical Programming*, 6:48–61, 1974.

[5] M. J. Osborne and A. Rubinstein. *A Course in Game Theory.* The MIT Press, Cambridge, Massachusetts, 1994.

[6] L. S. Shapley. On Balanced Sets and Cores. *Naval Research Logistics Quarterly*, 14(4):453–460, 1967.

[7] V. V. Vazirani. *Approximation Algorithms.* Springer–Verlag, Berlin/Heidelberg, 2001.