

Linear optimization and mathematical finance

Geir Dahl

Kristina Rognlien Dahl

Department of Mathematics, University of Oslo.

1 Introduction

A goal of this short note is to explain how linear optimization, also called linear programming (LP), may be used to solve two basic problems in mathematical finance concerning arbitrage and dominant trading strategies, respectively. The note is an extension of a previous version [1] where we want to make the presentation self-contained, with complete proofs of the duality theorem of linear optimization.

We focus on the discrete one-step model in mathematical finance. What we obtain are new proofs based on LP duality, and an efficient computational approach for these problems. The idea here is to start with some “natural” LP problems describing the investors problem. This note supplements the presentation in [3]. (The proofs in [3], see (1.9) and (1.16), are related, but different.)

However, as many students of mathematical finance are not familiar with LP, a brief introduction to this topic is in order. In particular, a proof of the important linear programming duality theorem will be given. Our proof is based on convex analysis, in particular the separating hyperplane theorem and Farkas’ lemma. For an introduction to linear optimization, including efficient algorithms, we recommend [4] (here a constructive proof of the duality theorem based on the simplex algorithm is given).

We treat vectors as column vectors and O denotes the zero vector (or matrix).

2 A preparation: distances and convex sets

We start with some preparations concerning nearest points of sets.

Recall from the course MAT1110 that a set $S \subseteq \mathbb{R}^n$ is *closed* if it contains the limit point of each convergent sequence of points in S . For a nonempty, closed set $S \subseteq \mathbb{R}^n$ and a point $z \in \mathbb{R}^n$ we define $d_S(z) = \inf\{\|s - z\| : s \in S\}$ and call this the *distance* of x to S . We here use the Euclidean norm. We say that $s_0 \in S$ is a *nearest point* of S to z if $\|s_0 - z\| = d_S(z)$. One can show that a nearest point always exists. In fact, $d_S(z) = d_{S'}(z)$ where $S' = \{s \in S : \|s - z\| \leq d_S + 1\}$ and the minimum distance from z to S' is attained as S' is closed and bounded (i.e., compact) and the Euclidean norm is continuous; recall here the extreme value theorem.

Thus, closedness of (a nonempty set) S assures that a nearest point of a given point exists. But such a point may not be unique. However, for a VERY interesting class of sets there is a unique nearest point! A set $C \subseteq \mathbb{R}^n$ is called *convex* if $\lambda x + (1 - \lambda)y \in C$ for all $x, y \in C$ and $\lambda \in [0, 1]$. Geometrically, this property means that whenever we choose two points in the set, say $x, y \in C$, then all points on the line segment between x and y also lie in C .

Lemma 1 (Unique nearest point for convex sets) *Let $C \subseteq \mathbb{R}^n$ be a non-empty closed convex set. Then, for every $z \in \mathbb{R}^n$, there is a unique nearest point c to z in C*

Proof. Assume that both c_0 and c_1 are nearest points to z in C , and let $d = d_C(z) = \|z - c_0\| = \|z - c_1\|$. Then c_0 and c_1 both lie on the boundary of the closed ball $B = \{y \in \mathbb{R}^n : \|y - z\| \leq d\}$ with radius d and center z . But the midpoint $c^* = (1/2)c_0 + (1/2)c_1$ lies in C , as C is convex, and c^* also lies in the interior of B . Therefore $\|c^* - z\| < d$, a contradiction. This proves that a nearest point must be unique. \square

3 The separating hyperplane theorem

A *hyperplane* is a "generalized plane". More formally, it is a set $H \subseteq \mathbb{R}^n$ of the form $H = \{x \in \mathbb{R}^n : a^T x = \alpha\}$ for some nonzero vector a and a real number α . The vector a is a *normal vector* of the hyperplane. In \mathbb{R}^2 , a hyperplane is a line, in \mathbb{R}^3 it is a plane. We denote $H_{a,\alpha} = H = \{x \in \mathbb{R}^n :$

$a^T x = \alpha$. Also, we define the *halfspaces*

$$\begin{aligned} H_{a,\alpha}^- &:= \{x \in \mathbb{R}^n : a^T x \leq \alpha\}; \\ H_{a,\alpha}^+ &:= \{x \in \mathbb{R}^n : a^T x \geq \alpha\}. \end{aligned} \tag{1}$$

These halfspaces represent the two sides of the hyperplane as we can convince ourselves in \mathbb{R}^2 or \mathbb{R}^3 .

Now, consider two sets S and T in \mathbb{R}^n . We say that the hyperplane $H_{a,\alpha}$ *strongly separates* S and T if there is an $\epsilon > 0$ such that $S \subseteq H_{a,\alpha-\epsilon}^-$ and $T \subseteq H_{a,\alpha+\epsilon}^+$ or vice versa. In \mathbb{R}^2 , this means that the sets S and T are on separate sides of the line $H_{a,\alpha}$, and neither of the sets intersect the line.

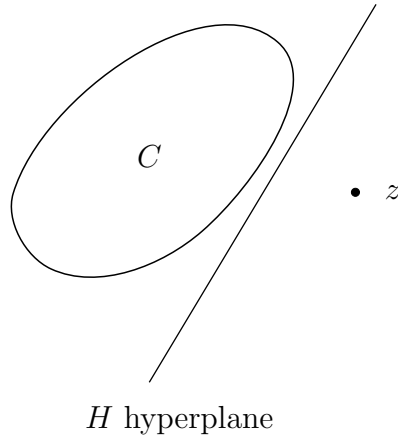


Figure 1: Strong separation

Theorem 2 (Separating hyperplane theorem) *Let $C \subseteq \mathbb{R}^n$ be a nonempty closed convex set. Let $z \in \mathbb{R}^n$ and assume that $z \notin C$. Then C and z can be strongly separated.*

Proof. Consider C and z as indicated above, and let p be the unique nearest point to x in C (see Lemma 1). Let $x \in C$ and let $0 < \lambda < 1$. Since C is convex, $(1 - \lambda)p + \lambda x \in C$ and since p is a nearest point we have that $\|(1 - \lambda)p + \lambda x - z\| \geq \|p - z\|$, i.e., $\|(p - z) + \lambda(x - p)\| \geq \|p - z\|$. By squaring both sides and calculating the inner products we obtain $\|p - z\|^2 + 2\lambda(p - z)^T(x - p) + \lambda^2\|x - p\|^2 \geq \|p - z\|^2$. We now subtract

$\|p - z\|^2$ on both sides, divide by λ , let $\lambda \rightarrow 0^+$ and finally multiply by -1 . This gives the inequality

$$(z - p)^T(x - p) \leq 0 \quad \text{for all } x \in C. \quad (2)$$

Consider the hyperplane H containing p and having normal vector $a := z - p$, i.e., $H = \{x \in \mathbb{R}^n : a^T x = \alpha\}$ where $\alpha = a^T p$. The inequality (2) shows that $C \subseteq H_{a,\alpha}^-$. Moreover, $z \notin H_{a,\alpha}^-$ as $z \neq p$ (because $z \notin C$). Now, consider the hyperplane H^* which is parallel to H (i.e., having the same normal vector) and contains the point $(1/2)(z + p)$. Then it is easy to see that H^* strongly separates z and C as desired. \square

An illustration is in Figure 1. There are several separation theorems for convex sets, see e.g. [2] and the references given there.

4 Farkas' lemma

Using the separating hyperplane theorem (Theorem 2) we will prove *Farkas' lemma* which characterizes when a system of linear inequalities has a solution. It is central in optimization, as we shall see in the next section.

Theorem 3 (Farkas' lemma) *Let A be a real $m \times n$ matrix and let $b \in \mathbb{R}^m$. Then there exists an $x \geq 0$ satisfying $Ax = b$ if and only if for each $y \in \mathbb{R}^m$ with $y^T A \geq 0$ it also holds that $y^T b \geq 0$.*

Proof. Let a^1, a^2, \dots, a^n denote the column vectors of the matrix A . Consider the set $C = \{\sum_{j=1}^n \lambda_j a^j : \lambda_j \geq 0 \text{ for } j = 1, \dots, n\} \subseteq \mathbb{R}^m$; this set is called the *convex cone generated by a^1, a^2, \dots, a^n* , see [2]. Then C is closed (this can be checked from the definition of C , but it also follows from Proposition 2.5.5 in [2]). We observe that $Ax = b$ has a nonnegative solution x if and only if $b \in C$.

Assume now that x satisfies $Ax = b$ and $x \geq 0$. If $y^T A \geq 0$, then $y^T b = y^T(Ax) = (y^T A)x \geq 0$ as the inner product of two nonnegative vectors. Conversely, if $Ax = b$ has no nonnegative solution, then $b \notin C$. But then, by the separating hyperplane theorem (Theorem 2), C and b can be strongly separated, so there is a nonzero vector $y \in \mathbb{R}^m$ and $\alpha \in \mathbb{R}$ with $y^T x \geq \alpha$ for each $x \in C$ and $y^T b < \alpha$. As $0 \in C$, we have $\alpha \leq 0$. We claim that $y^T x \geq 0$ for each $x \in C$: for if $y^T x < 0$ for some $x \in C$, there would be a point $\lambda x \in C$ with $\lambda > 0$ such that $y^T(\lambda x) < \alpha$, a contradiction. Therefore

(as $a^j \in C$) $y^T a^j \geq 0$ so $y^T A \geq O$. Since $y^T b < 0$ we have proved the other direction of Farkas' lemma. \square

Farkas' lemma can be understood geometrically: b lies in the cone C defined in the proof above if and only if there is no hyperplane $H = \{x \in \mathbb{R}^n : y^T x = 0\}$ (having normal vector y) that separates b and C , i.e., $y^T b < 0$ and $y^T a^j \geq 0$ for each j .

5 Linear programming duality

Linear programming (LP), or linear optimization, is to maximize a linear function in n variables subject to a finite number of linear constraints. These constraints are linear equations and/or linear inequalities. A recommended book in LP is [4]. Hence, a standard linear programming problem is of the form

$$\sup\{c^T x : Ax \leq b\} \tag{3}$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and the $m \times n$ matrix A are given, and the variable vector is $x \in \mathbb{R}^n$. Here, \leq means componentwise inequality (i.e., it holds for each component).

We call problem (3) the *primal* problem. The primal problem is *feasible* if there is an x satisfying $Ax \leq b$, and such an x is called a *feasible solution*. A feasible solution x_0 is *optimal* if $c^T x_0 = \sup\{c^T x : Ax \leq b\}$, so then the supremum is attained and we may write max in stead of sup. We define the supremum in (3) to be $-\infty$ if the problem is not feasible, and it is $+\infty$ if the problem is *unbounded*, meaning that there is a sequence (x_k) of feasible solutions such that $c^T x_k \rightarrow \infty$ as $k \rightarrow \infty$.

Associated to each (primal) LP problem there is another LP problem, called its *dual* problem. The dual problem associated with problem (3) is

$$\inf\{b^T y : A^T y = c, y \geq O\}. \tag{4}$$

As for the primal problem we use the terms feasible problem, feasible solution and optimal solution for the dual problem. We define the infimum in (4) to be ∞ if the problem is not feasible, and it is $-\infty$ if the problem is *unbounded*, i.e., there is a sequence (y_k) of feasible solutions such that $b^T y_k \rightarrow -\infty$ as $k \rightarrow \infty$.

One of the main theorems in optimization is the LP duality theorem (see [2], [4]).

Theorem 4 (Linear programming duality theorem) (i) Assume that the primal problem (3) has an optimal solution. Then the dual problem (4) also has an optimal solution and

$$\max\{c^T x : Ax \leq b\} = \min\{b^T y : A^T y = c, y \geq O\}. \quad (5)$$

(ii) If one of the problems is unbounded, then the other problem is not feasible. Thus, when at least one problem is feasible, $\sup\{c^T x : Ax \leq b\} = \inf\{b^T y : A^T y = c, y \geq O\}$.

Proof. (i) Let x be feasible in the primal problem and y feasible in the dual problem, so $Ax \leq b$ and $A^T y = c, y \geq O$. Then

$$c^T x = (A^T y)^T x = y^T Ax \leq y^T b = b^T y$$

where the inequality follows from $Ax \leq b$ as $y \geq O$. By first taking the supremum over feasible x and then the infimum over feasible y in this inequality we obtain

$$(*) \quad \sup\{c^T x : Ax \leq b\} \leq \inf\{b^T y : A^T y = c, y \geq O\}.$$

Let x_0 be an optimal solution of the primal problem. Let a_i^T denote the i th row in the matrix A (so a_i is a column vector). Define $I = \{i \leq m : a_i^T x_0 = b_i\}$ which corresponds to the indices of inequalities from $Ax \leq b$ that hold with equality for $x = x_0$.

Claim: For each $z \in \mathbb{R}^n$ satisfying $a_i^T z \leq 0$ for all $i \in I$, the inequality $c^T z \leq 0$ also holds.

Otherwise, there is a $z \in \mathbb{R}^n$ with $a_i^T z \leq 0$ for all $i \in I$ and $c^T z > 0$. Then, for suitably small $\epsilon > 0$, the point $x' = x_0 + \epsilon z$ satisfies $Ax' \leq b$ because (i) for each $i \in I$ we have $a_i^T x' = a_i^T x_0 + \epsilon a_i^T z = b_i + \epsilon a_i^T z \leq b_i$, and (ii) for each $i \leq m$ with $i \notin I$ we have $a_i^T x' = a_i^T x_0 + \epsilon a_i^T z < b_i + \epsilon a_i^T z$, so $a_i^T x' \leq b_i$ for ϵ small. But $c^T x' = c^T x_0 + \epsilon c^T z > c^T x_0$ which contradicts that x_0 is an optimal solution. This proves the Claim.

Next, the Claim makes it possible to apply Farkas' lemma (Theorem 3) to the matrix A' whose columns are the vectors a_i for $i \in I$ (so A' plays the role of A in Theorem 3). As a result there must exist nonnegative numbers y_i for $i \in I$ such that $\sum_{i \in I} y_i a_i = c$. Therefore $A^T y = c$ where $y \in \mathbb{R}^m$ is the vector with components y_i for $i \in I$ (those we just found), and $y_i = 0$

otherwise. Since $A^T y = c$ and $y \geq O$, y is a feasible solution in the dual problem. Moreover, using that $y_i = 0$ for $i \notin I$ we get

$$c^T x_0 = y^T A x_0 = \sum_{i \in I} y_i (a_i^T x_0) = \sum_{i \in I} y_i b_i = b^T y.$$

This proves, due to the inequality (*), that y is an optimal solution of the dual problem and that the maximum in the primal problem equals the minimum in the dual problem, and (5) holds.

(ii) Consider the inequality (*). If the primal problem is unbounded, then the dual problem is not feasible (for, due to (*), $b^T y$ would be an upper bound on $c^T x$). So, in this case, both sides of (*) are ∞ . Similarly, if the dual problem is unbounded, both sides of (*) are $-\infty$. Finally, if the supremum in (*) is finite, one can show that the supremum is attained and therefore, by the first part of the theorem, "sup=max=min=inf". The same is true when the infimum is finite. (We omit the detailed argument here, it involves the structure of polyhedra - the feasible sets of LP problems). \square

Thus, for an LP problem, there are only three possible situations: (i) it is not feasible (i.e., no feasible solution exist), (ii) it is unbounded, or (iii) an optimal solution exists; then both problems have an optimal solution, and the corresponding optimal values are equal.

Note the special case where $b = O$ (in our primal and dual problems): then the function to be minimized in the dual is constant equal to 0. Thus, the problem is simply to determine if there are any feasible solutions in the dual. This special case will be useful later.

6 The fundamental theorem of asset pricing via linear programming

In this section, we will apply LP theory in order to prove a version of the fundamental theorem of asset pricing and also to find dominant trading strategies. See [3] for a discussion of these notions and mathematical finance.

We shall use the following notation:

- K : number of states (scenarios), n : number of assets
- $P = [p_{ij}]$: payoff matrix of size $K \times n$ where p_{ij} is payoff under state i for asset j (this is $\Delta S_j^*(\omega_i)$ in Pliska's notation)

- $h \in \mathbb{R}^n$: a trading strategy, we buy h_j units of asset j
- $x \in \mathbb{R}^K$: the payoff, i.e., outcome of some trading strategy under different states
- O, e, I : the zero vector (or zero matrix) O , the all ones vector e and the identity matrix I (of suitable size). We treat vectors as column vectors.
- $\text{Nul}(A), \text{Col}(A)$: nullspace and columnspace of a matrix A

A *risk-neutral probability measure* is a vector y with positive components that sum to 1 such that the dot product of y and each column of P is zero (meaning that expected payoff of each asset is zero).

The first LP model: find an arbitrage

Consider the LP problem

$$\begin{aligned}
 & \max && \sum_{i=1}^K x_i \\
 & \text{subject to} && \\
 & && x = Ph \\
 & && x \geq O
 \end{aligned} \tag{6}$$

Here $x = Ph$ relates payoff x and trading strategy h ; the linear equation says that x is a linear combination of the columns in P . The nonnegativity of x is very desirable: we will not lose money under any state. The objective (goal) is to maximize the sum of the payoffs, where we sum over all states. It should be a reasonable goal and can (if you like) be given a probabilistic interpretation. The main point is that it reflects that we look for positive payoffs for at least one scenario, i.e., an arbitrage possibility.

So: *an arbitrage exists if and only if the optimal value of the LP problem (6) is positive*. Since LP problems can be solved very fast using different algorithms (e.g. the simplex algorithm), we can find an arbitrage, or prove that it does not exist, efficiently, even if K and n are very large. For instance, even with some thousands of assets and several hundreds of states it should only take a couple of seconds to solve the problem assuming you have a good LP code. Moreover, we may obtain an important theoretical result from this LP viewpoint.

Theorem 5 (*The arbitrage theorem*) *There is no arbitrage if and only if there is a risk-neutral probability measure.*

Proof. We prove the arbitrage theorem by applying duality theory to our LP above, and doing some matrix calculations for partitioned matrices.

The LP problem (6) may be written in the form of the primal problem in (3) as follows:

$$\max \left\{ \begin{bmatrix} O \\ e \end{bmatrix}^T \begin{bmatrix} h \\ x \end{bmatrix} : \begin{bmatrix} P & -I \\ -P & I \\ O & -I \end{bmatrix} \begin{bmatrix} h \\ x \end{bmatrix} \leq \begin{bmatrix} O \\ O \\ O \end{bmatrix} \right\}$$

Here we used that $Ph = x$ is equivalent to $Ph - x \leq O$, $-Ph + x \leq O$. The dual of this problem (confer (4)) is

$$\min \left\{ \begin{bmatrix} O \\ O \\ O \end{bmatrix}^T \begin{bmatrix} y^1 \\ y^2 \\ y^3 \end{bmatrix} : \begin{bmatrix} P^T & -P^T & O \\ -I & I & -I \end{bmatrix} \begin{bmatrix} y^1 \\ y^2 \\ y^3 \end{bmatrix} = \begin{bmatrix} O \\ e \end{bmatrix}, y^1, y^2, y^3 \geq O \right\}$$

Note that the objective function is simply 0! By using the substitution $y = y^2 - y^1$, $z = y^3$ the dual problem becomes

$$\min \{0 : P^T y = O, y = z + e, z \geq O\}.$$

Check this! This problem has a feasible solution y if and only if there is a vector $y \in \text{Nul}(P^T) = \text{Col}(P)^\perp$ such that $y \geq e$. This, again, must be equivalent to the existence of an $y \in \text{Nul}(P^T)$ with $y_i > 0$ ($i \leq K$) and $\sum_i y_i = 1$; this follows by suitable scaling of y . Summing up, we have verified that the following statements are equivalent:

- there is no arbitrage
- the optimal value in the LP problem (6) is zero
- there is a strictly positive vector $y \in \text{Nul}(P^T)$ with $\sum_i y_i = 1$; this is precisely a risk-neutral probability measure.

So the proof is complete. □

We now turn to the second theorem; it concerns dominant trading strategies.

The second LP model: find a dominant trading strategy

This problem is very similar to (6) but it contains an extra variable $\epsilon \in \mathbb{R}$:

$$\begin{aligned} \max \quad & \epsilon \\ \text{subject to} \quad & \\ & x = Ph \\ & x \geq \epsilon e \end{aligned} \tag{7}$$

The final constraints means that $x_j \geq \epsilon$ for each $j \leq n$, and the goal is to find a trading strategy which maximizes the minimum outcome (the optimal ϵ)! Note that (7) has feasible solutions (e.g., the zero vector).

So: *a dominant trading strategy exists if and only if the optimal value of the LP problem (7) is positive.*

Recall that a *linear pricing measure* is just like a risk-neutral probability measure, except that some probabilities may be zero.

Theorem 6 (*Dominant trading strategy/linear pricing measure*) *There is no dominant trading strategy if and only if there is a linear pricing measure.*

Proof. The proof is very similar to the previous one. First, we write the LP problem (7) in the form of the primal problem in (*):

$$\max \left\{ \begin{bmatrix} O \\ O \\ 1 \end{bmatrix}^T \begin{bmatrix} h \\ x \\ \epsilon \end{bmatrix} : \begin{bmatrix} P & -I & O \\ -P & I & O \\ O & -I & e \end{bmatrix} \begin{bmatrix} h \\ x \\ \epsilon \end{bmatrix} \leq \begin{bmatrix} O \\ O \\ O \end{bmatrix} \right\}$$

The dual of this problem is (see (*) and do a calculation as before):

$$\min\{0 : P^T(y^1 - y^2) = 0, -y^1 + y^2 - y^3 = 0, e^T y^3 = 1, y^1, y^2, y^3 \geq O\}.$$

With the substitution $y = y^2 - y^1$ and $\pi = y^3$ we see that we can eliminate y , and the problem simplifies to

$$\min\{0 : P^T \pi = 0, \sum_j \pi_j = 1, \pi \geq O\}.$$

A feasible solution in this problem is precisely a linear pricing measure. This proves that the following statements are equivalent:

- there exists a linear pricing measure
- the optimal value in the LP problem (7) is zero
- there is no dominant trading strategy

That's it!

□

Final comments:

1. Using LP problems (6) and (7) you may solve efficiently arbitrage and dominant trading strategy problems. Note that any (good) LP solver solves both the primal and the dual. So, e.g., if a linear pricing measure exist, you will get it!
2. The LP approach above may be extended in different ways. For instance, in (6), you may add the constraint $h \geq 0$ (no short-selling), and add an upper bound on the components of h , or further linear constraints on the payoff vector x . Well, the problem then gets more complex, but it is still an LP problem, and should be easy to solve computationally.
3. Perhaps this motivates you to learn more about LP: take a look at the course page for INF-MAT3370 Linear optimization <http://www.uio.no/studier/emner/matnat/ifi/INF-MAT3370/> There is also a Master course (INF-MAT5360) in optimization and convexity (see [2]), both useful areas for mathematical finance.

References

- [1] G. Dahl. Mathematical finance and linear programming (optimization). Notes, University of Oslo, 2009.
- [2] G. Dahl. An Introduction to Convexity. Lecture notes, University of Oslo, 2000 (see <http://folk.uio.no/geird/>)
- [3] S.R. Pliska. *Introduction to Mathematical Finance: Discrete Time Models*. Blackwell Publishers, 1997.
- [4] R. Vanderbei. *Linear programming: foundations and extensions*. Springer, Third edition, 2008.