A separating hyperplane theorem, the fundamental theorem of asset pricing, and Markov’s principle

Josef Berger∗ Gregor Svindland

July 15, 2015

Abstract

We prove constructively that every uniformly continuous convex function \( f : X \to \mathbb{R}^+ \) has positive infimum, where \( X \) is the convex hull of finitely many vectors. Using this result, we prove that a separating hyperplane theorem, the fundamental theorem of asset pricing, and Markov’s principle are constructively equivalent. This is the first time that important theorems are classified into Markov’s principle within constructive reverse mathematics.

Keywords: Separating hyperplane theorem, the fundamental theorem of asset pricing, Markov’s principle, constructive mathematics, reverse mathematics.

MSC2010 classification: 03B30, 03F60.

Constructive mathematics in the tradition of Errett Bishop [2, 3] is characterised by not using the law of excluded middle as a proof tool. As a major consequence, properties of the real number line like the limited principle of omniscience

\[ \text{LPO} \quad \forall x, y \in \mathbb{R} \ (x < y \lor x > y \lor x = y), \]

the lesser limited principle of omniscience

\[ \text{LLPO} \quad \forall x, y \in \mathbb{R} \ (x \leq y \lor x \geq y), \]

and Markov’s principle

\[ \text{MP} \quad \forall x \in \mathbb{R} \ (\neg (x = 0) \Rightarrow |x| > 0) \]

∗corresponding author
are no longer provable propositions but rather considered additional axioms. In this context, reverse mathematics attempts on deciding which axioms of this kind are necessary and sufficient to prove certain theorems. For example, the Hahn-Banach theorem is equivalent to LLPO [8]. Many properties of the reals still hold constructively:

- \( x \geq y \iff \neg (x < y) \)
- \( x = y \iff \neg \neg (x = y) \)
- \( |x| \cdot |y| > 0 \iff |x| > 0 \& |y| > 0 \)
- \( |x| > 0 \iff x > 0 \lor x < 0 \)

We cannot prove constructively that every nonempty bounded set of reals has an infimum. This gives rise to the following definition. An inhabited\(^1\) subset \( \mathcal{Y} \) of a metric space \( (X, d) \) is located if

\[
d(x, \mathcal{Y}) = \inf \{d(x, y) \mid y \in \mathcal{Y}\}
\]

exists for all \( x \in X \). If \( \mathcal{Y} \) is located and bounded and \( f : \mathcal{Y} \to \mathbb{R} \) is uniformly continuous, then

\[
\inf f = \inf \{f(y) \mid y \in \mathcal{Y}\}
\]

does exist [5, Corollary 2.2.7]. In this context, Brouwer’s fan theorem can be stated as follows [9].

**FAN** If \( f : [0, 1] \to \mathbb{R}^+ \) is uniformly continuous, then \( \inf f > 0 \).

Note that FAN can be deduced from LLPO [7]. There are many propositions which are equivalent to principles like LLPO or the fan theorem. In this paper, for the first time, we prove that important classical theorems are classified into Markov’s principle.

Set

\[
\mathcal{Y}_n = \left\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid \sum_{i=1}^n x_i > 0 \text{ and } 0 \leq x_i \text{ for all } i \right\},
\]

\[
\mathcal{X}_n = \left\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 1 \text{ and } 0 \leq x_i \text{ for all } i \right\},
\]

\(^1\)A set is inhabited if it contains an element, which is classically equivalent to being nonempty.
The following results on convexity, which are used for the announced classification later on, are interesting on their own.

Let \( C \) be an inhabited convex set. A function \( f : C \to \mathbb{R} \) is convex if

\[
f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)
\]

for all \( \lambda \in [0, 1] \) and \( x, y \in C \).

The following Proposition can be considered a constructive version of the fan theorem.

**Proposition 1.** Every uniformly continuous convex function \( f : X_m \to \mathbb{R}^+ \) has positive infimum, i.e. \( \inf f > 0 \).

**Proof.** We use induction on \( m \) to show the following statement: fix elements \( A_1, \ldots, A_m \) of a normed space \( V \) and a positive valued uniformly continuous convex function \( f \) on

\[
C(A_1, \ldots, A_m) = \left\{ \sum_{i=1}^{m} \lambda_i A_i \mid \lambda \in X_m \right\}.
\]

Then \( f \) has positive infimum.

For \( m = 1 \), the statement is trivial. Suppose that the statement holds for \( m \geq 1 \). Fix \( A_1, \ldots, A_{m+1} \in V \) and a uniformly continuous convex function \( f : C(A_1, \ldots, A_{m+1}) \to \mathbb{R}^+ \). Note that \( C(A_1, \ldots, A_{m+1}) \) is located and bounded and set \( \alpha = \inf f \). We define a decreasing sequence of subintervals \( [a_n, b_n] \) of \([0, 1]\) such that

- \( b_n - a_n \in \{0, 2^{-n}\} \)
- \( a_n = b_n \Rightarrow 0 < \alpha \)
- \( 0 < \alpha \lor \alpha = \inf \{ f(P) \mid P \in C \} \)

for every \( n \in \mathbb{N} \), where

\[
C_n = C(A_1 + a_n(A_2 - A_1), A_1 + b_n(A_2 - A_1), A_3, \ldots, A_{m+1}).
\]

Set \( a_0 = 0 \) and \( b_0 = 1 \) and suppose that \( a_n \) and \( b_n \) have been defined.
case 1  If \( a_n = b_n \), set \( a_{n+1} = b_{n+1} = a_n \).

case 2  Suppose that \( a_n < b_n \). Set \( c = \frac{a_n + b_n}{2} \).

case 2.1  If \( 0 < \alpha \), set \( a_{n+1} = b_{n+1} = c \).

case 2.2  Suppose that \( \alpha = \inf \{ f(P) \mid P \in C_n \} \). Set

\[
\varepsilon = \frac{1}{2} \inf \{ f(P) \mid P \in C(A_1 + c(A_2 - A_1), A_3, \ldots, A_{m+1}) \}.
\]

By the induction hypothesis, \( \varepsilon > 0 \).

case 2.2.1  If \( 0 < \alpha \), set \( a_{n+1} = b_{n+1} = c \).

case 2.2.2  If \( \alpha \leq \varepsilon \), there is

\[
X = \lambda_1(A_1 + a_n(A_2 - A_1)) + \lambda_2(A_1 + b_n(A_2 - A_1)) + \sum_{i=3}^{m+1} \lambda_i A_i \in C_n
\]

with

\[
f(X) < \varepsilon.
\] (2)

We show that

\[
d(X, C(A_1 + c(A_2 - A_1), A_3, \ldots, A_{m+1})) > 0.
\] (3)

Suppose that

\[
d(X, C(A_1 + c(A_2 - A_1), A_3, \ldots, A_{m+1})) < \delta,
\]

for \( \delta > 0 \) with

\[
\|P - P'\| < \delta \quad \Rightarrow \quad |f(P) - f(P')| < \varepsilon \quad (P, P' \in C(A_1, \ldots, A_{m+1})).
\]

Then there is \( P \in C(A_1 + c(A_2 - A_1), A_3, \ldots, A_{m+1}) \) such that \( \|X - P\| < \delta \) and hence

\[
f(X) > f(P) - \varepsilon \geq 2\varepsilon - \varepsilon = \varepsilon,
\]

in contradiction to (2). This completes the proof of (3), which implies that we have either \( \lambda_1 > \lambda_2 \) or \( \lambda_1 < \lambda_2 \).

case 2.2.2.1  If \( \lambda_1 > \lambda_2 \), then for any

\[
P = \mu_1(A_1 + a_n(A_2 - A_1)) + \mu_2(A_1 + b_n(A_2 - A_1)) + \sum_{i=3}^{m+1} \mu_i A_i \in C_n
\]

such that \( \mu_1 < \mu_2 \) there is \( \nu \in (0, 1) \) such that \( \nu(\lambda_1 - \lambda_2) + (1-\nu)(\mu_1 - \mu_2) = 0 \), and hence

\[
\nu X + (1 - \nu) P \in C(A_1 + c(A_2 - A_1), A_3, \ldots, A_{m+1}).
\]
We therefore obtain
\[ 2\varepsilon \leq f(\nu X + (1 - \nu)P) \leq \nu f(X) + (1 - \nu) f(P) < \nu \varepsilon + (1 - \nu) f(P), \]
which implies that \( f(P) > \varepsilon \). Hence, points like \( P \) are irrelevant for the calculation of \( \alpha \), and we may proceed by setting \( a_{n+1} = a_n \) and \( b_{n+1} = c \).

**case 2.2.2.2** Similarly, if \( \lambda_1 < \lambda_2 \), set \( a_{n+1} = c \) and \( b_{n+1} = b_n \).

Let \( a \) be the limit of \( (a_n) \) (and thus also of \( (b_n) \)). If \( 0 < \alpha \), we are done. If \( \alpha < \inf \{ f(P) \mid P \in C(A_1 + a(A_2 - A_1), A_3, \ldots, A_{m+1}) \} \), by the uniform continuity of \( f \), we can find an \( n \) such that
\[ \alpha < \inf \{ f(P) \mid P \in C_n \}, \]
thus \( 0 < \alpha \).

Note that the proof of Proposition 1 is given for normed spaces. Indeed, Proposition 1 is equivalent to the following seemingly stronger result:

**Corollary 1.** Suppose that \( f : C(A_1, \ldots, A_m) \to \mathbb{R}^+ \) is convex and uniformly continuous, where \( A_1, \ldots, A_m \) are points in a normed space, and \( C(A_1, \ldots, A_m) \) is given by (1). Then \( f \) has positive infimum.

The equivalence of Proposition 1 and Corollary 1 follows from the fact that for any function \( f \) as in Corollary 1 the composition \( f \circ \kappa \) satisfies the requirements of Proposition 1, where
\[ \kappa : \mathcal{X}_m \to C(A_1, \ldots, A_m), (\lambda_1, \ldots, \lambda_m) \mapsto \sum_{i=1}^m \lambda_i A_i. \]

The **closure** \( \overline{\mathcal{Y}} \) of a set \( \mathcal{Y} \) is the set of all limits of sequences in \( \mathcal{Y} \).

**Lemma 1.** Let \( \mathcal{Y} \) be an inhabited convex subset of a Hilbert space \( (H, \langle \cdot, \cdot \rangle) \). Fix \( x \in H \) and assume that \( d = d(x, \mathcal{Y}) \) exists. Then there exists a unique \( a \in \overline{\mathcal{Y}} \) such that \( \|a - x\| = d \). Furthermore, we have
\[ \langle a - x, c - a \rangle \geq 0 \]
and therefore
\[ \langle a - x, c - x \rangle \geq d^2 \]
for all \( c \in \mathcal{Y} \).
Proof. Fix a sequence \((c_l)\) in \(Y\) such that \(\|c_l - x\| \to d\). Since
\[
\|c_m - c_l\|^2 = \|(c_m - x) - (c_l - x)\|^2 = \\
2\|c_m - x\|^2 + 2\|c_l - x\|^2 - 4\|\frac{c_m + c_l}{2} - x\|^2 \leq \\
2\left(\|c_m - x\|^2 - d^2\right) + 2\left(\|c_l - x\|^2 - d^2\right),
\]
\((c_l)\) is a Cauchy sequence and therefore converges to an \(a \in Y\). Since \(\|c_l - x\| \to \|a - x\|\), we obtain \(\|a - x\| = d\). Now fix \(b \in Y\) with \(\|b - x\| = d\). Then
\[
\|a - b\|^2 = \|(a - x) - (b - x)\|^2 = \\
2\|a - x\|^2 + 2\|b - x\|^2 - 4\|\frac{a + b}{2} - x\|^2 \leq 0,
\]
thus \(a = b\).

Fix \(c \in Y\) and \(\lambda \in (0,1)\). Since
\[
\|a - x\|^2 \leq \|(1 - \lambda)a + \lambda c - x\|^2 = \|(a - x) + \lambda(c - a)\|^2 = \\
\|a - x\|^2 + \lambda^2\|c - a\|^2 + 2\lambda\langle a - x, c - a \rangle,
\]
we obtain
\[
0 \leq \lambda\|c - a\|^2 + 2\langle a - x, c - a \rangle.
\]
Since \(\lambda\) can be arbitrarily small, we can conclude that
\[
\langle a - x, c - a \rangle \geq 0.
\]
This also implies that
\[
\langle a - x, c - x \rangle = \langle a - x, c - a \rangle + \langle a - x, a - x \rangle \geq d^2.
\]

Proposition 1 and Lemma 1 imply the following constructive separating hyperplane theorem; compare to [5, Theorem 5.2.9].

Corollary 2. Let \(C\) be a convex closed located subset of a Hilbert space \((H, \langle , \rangle)\). Moreover, let \(x_1, \ldots, x_n \in H\) such that \(d(x,c) > 0\) for all \(x \in C(x_1, \ldots, x_n)\) and \(c \in C\). Then there exist \(\varepsilon > 0\) and \(p \in H\) such that
\[
\langle p, x - c \rangle \geq \varepsilon
\]
for all \(x \in C(x_1, \ldots, x_n)\) and \(c \in C\).
Proof. Since
\[ |d(x, C) - d(y, C)| \leq d(x, y) \]
for all \( x, y \in H \), the function
\[ \kappa : C(x_1, \ldots, x_n) \to \mathbb{R}, x \mapsto d(x, C) \]
is uniformly continuous. Since \( C \) is closed, \( \kappa \) is positive valued, see Lemma 1. Since \( C \) is convex, \( \kappa \) is convex as well. By Corollary 1, \( \inf \kappa > 0 \). The set
\[ \mathcal{Y} = \{ x - c \mid x \in C(x_1, \ldots, x_n), c \in C \} \]
is convex. Note that \( d(0, \mathcal{Y}) = \inf \kappa \) and apply Lemma 1. \( \square \)

Vectors \( x_1, \ldots, x_n \) are \textit{linearly independent} if for all \( \lambda \in \mathbb{R}^n \) the implication
\[ \sum_{i=1}^n |\lambda_i| > 0 \Rightarrow \| \sum_{i=1}^n \lambda_i x_i \| > 0 \]
is valid. Such vectors span located subsets [5, Lemma 4.1.2]:

Lemma 2. If \( x_1, \ldots, x_m \in \mathbb{R}^n \) are linearly independent, then the set
\[ \left\{ \sum_{i=1}^m \xi_i x_i \mid \xi \in \mathbb{R}^m \right\} \]
is convex, closed, and located.

Now we are ready for the fundamental theorem of asset pricing. Consider a portfolio of \( m \) stocks and one bond. The prices of the assets at time 0 (=present) are known and denoted by \( \pi \in \mathbb{R}^{m+1} \). The prices of the stocks at time 1 (=future) are unknown, but we know that there are \( n \) possible developments and we know the prices for each case. All prices are positive and the price of the bond is always 1, also at time 1.\(^2\) Let \( c_{ij} \) denote the price of the \( i \)-th asset in case \( j \) and set \( C = (c_{ij}) \). We assume that the first coordinate corresponds to the bond, i.e. \( \pi_1 = 1 \) and also \( c_{1j} = 1 \) for all \( j = 1, \ldots, n \).

A vector \( p \in \mathcal{P}_n \) is called \textit{equivalent martingale measure} if \( C \cdot p = \pi \).

A vector \( \xi \in \mathbb{R}^{m+1} \) is called \textit{arbitrage strategy} if
\[ \xi \cdot \pi \leq 0 \quad \text{and} \quad \xi \cdot C \in \mathcal{Y}_n. \]

\(^2\)This means that the prices are discounted.
Note that every $\xi \in \mathbb{R}^{m+1}$ corresponds to a trading strategy, where $\xi_i, i = 1, \ldots, m + 1$, denotes the number of shares of asset $i$ that the trader buys. Hence, the price of $\xi$ is $\xi \cdot \pi$ and the payoff at time 1 over all possible future scenarios is $\xi \cdot C$. Thus arbitrage strategies are trading strategies which correspond to riskless gains, since they do not cost anything ($\xi \cdot \pi \leq 0$) and do not produce any losses, and even a strict gain for at least one possible future scenario ($\xi \cdot C \in \mathcal{Y}_n$).

**Lemma 3.** The following are equivalent:

1) There exists an arbitrage strategy $\xi$.

2) There exists a vector $\mu \in \mathbb{R}^m$ such that

$$
\mu \cdot (C - \pi) \in \mathcal{Y}_n,
$$

where $C - \pi := (c_{ij} - \pi_i)_{i=2,\ldots,m+1;j=1,\ldots,n}$.

**Proof.** 1) $\Rightarrow$ 2) Just set $\mu = (\xi_2, \ldots, \xi_{m+1})$.

2) $\Rightarrow$ 1) Set $\xi := (-\sum_{i=1}^m \mu_i \pi_{i+1}, \mu_1, \ldots, \mu_m)$. Since $\pi_1 = 1$, we obtain

$$
\xi \cdot \pi = -\sum_{i=1}^m \mu_i \pi_{i+1} + \sum_{i=1}^m \mu_i \pi_{i+1} = 0.
$$

Since $c_{1j} = 1$, we obtain

$$
(\xi \cdot C)_j = \sum_{i=1}^m \mu_i (c_{(i+1)j} - \pi_{i+1}) = (\mu \cdot (C - \pi))_j.
$$

Thus $\xi \cdot C \in \mathcal{Y}_n$.

The classically provable fundamental theorem of asset pricing states that there exists an equivalent martingale measure if and only if there doesn’t exist an arbitrage strategy. This is probably the most fundamental theorem of mathematical finance and means that in ‘fair’ markets, i.e. in arbitragefree markets, the ‘fair’ prices are given by expectations under equivalent martingale measures. There is a vast literature on this result; we refer to [6] for a comprehensive discussion of the fundamental theorem of asset pricing.

Let $A := C - \pi$. By Lemma 3, excluding the existence of arbitrage strategies is the same as claiming that

$$
\{\xi \cdot A \mid \xi \in \mathbb{R}^m\} \cap \mathcal{Y}_n = \emptyset.
$$
Hence, the fundamental theorem of asset pricing reads as

\[ \{ \xi \cdot A \mid \xi \in \mathbb{R}^m \} \cap \mathcal{Y}_n = \emptyset \iff \exists p \in \mathcal{P}_n (A \cdot p = 0). \] (4)

The implication ‘\(\Leftarrow\)’ in (4) is easily verified, since for any \(\xi \in \mathbb{R}^m\) it follows that

\[ (\xi \cdot A) \cdot p = \xi \cdot (A \cdot p) = 0, \]

which implies that \(\neg(\xi \cdot A \in \mathcal{Y}_n)\). The non-trivial implication in (4) is ‘\(\Rightarrow\)’. In what follows, for technical reasons, we will assume that the rows of the matrix \(A\) are linearly independent. This is a standard assumption even classically, since it simply means that there are no redundant stocks in the market, i.e. stocks that can be hedged by others. Then the non-trivial implication of (4) reads as follows:

**FTAP** For every \(\mathbb{R}^{m \times n}\)-matrix \(A\) with linearly independent rows we have

\[ \{ \xi \cdot A \mid \xi \in \mathbb{R}^m \} \cap \mathcal{Y}_n = \emptyset \Rightarrow \exists p \in \mathcal{P}_n (A \cdot p = 0). \]

We will show that FTAP is equivalent to Markov’s principle and indeed also equivalent to the following version of a classically well-known separation result:

**SEP** Let \((H, \langle \cdot, \cdot \rangle)\) be a Hilbert space and suppose that \(C \subseteq H\) is convex, closed, and located. Moreover, let \(x_1, \ldots, x_n \in H\) and suppose that \(C \cap C(x_1, \ldots, x_n) = \emptyset\).

Then there exists \(\varepsilon > 0\) and a \(p \in H\) such that

\[ \langle p, x - c \rangle \geq \varepsilon \]

for all \(x \in C(x_1, \ldots, x_n)\) and \(c \in C\).

The classical version of SEP says that any two nonempty closed convex sets may be strictly separated by a hyperplane, provided that their intersection is void and one of the sets is compact; see for instance [1] Theorem 5.79.

**Proposition 2.** The following are equivalent:

a) SEP

b) FTAP

c) MP

9
Proof. a) ⇒ b): Assume SEP. Fix a $\mathbb{R}^{m \times n}$-matrix $A$ with linearly independent rows. By Lemma 2, the set
\[ C = \{ \xi \cdot A \mid \xi \in \mathbb{R}^m \} \]
is convex, closed, and located. Assume that
\[ C \cap X_n = \emptyset. \]
Then in particular $C \cap X_n = \emptyset$, and by SEP we obtain
\[ \exists \varepsilon > 0 \exists q \in \mathbb{R}^n \forall x \in X_n, c \in C \left( \langle q, x \rangle \geq \langle q, c \rangle + \varepsilon \right). \]
Let $e_i$ be the $i$-th unit vector in $\mathbb{R}^n$. Since $0 \in C$, we obtain $\langle q, e_i \rangle \geq \varepsilon$. This shows that all components of $q$ are positive. Now fix $c \in C$ and assume that $|\langle q, c \rangle| > 0$. Since $C$ is a linear space, there exists $d \in C$ such that
\[ \langle q, e_i \rangle < \langle q, d \rangle + \varepsilon, \]
a contradiction which shows that $A \cdot q = 0$. Therefore,
\[ p = \left( \frac{q_1}{\sum_{i=1}^n q_i}, \ldots, \frac{q_n}{\sum_{i=1}^n q_i} \right) \]
is an element of $P_n$ with $A \cdot p = 0$.

b) ⇒ c): Now assume FTAP and fix a real number $a$ with $-a = 0$. Consider the Matrix $A = (|a|, -1)$. Suppose that there exists a $\xi \in \mathbb{R}$ with $(\xi |a|, -\xi) \in Y_2$. In the first case, we have $\xi |a| > 0$ and $-\xi \geq 0$ and therefore $\xi > 0$ and $\xi \leq 0$. In the second case, we have $\xi |a| \geq 0$ and $-\xi > 0$ and therefore $a = 0$. So there cannot be such a $\xi$. Now FTAP yields the existence of a $p \in P_2$ with $p_1 |a| = p_2$. This implies that $|a| > 0$.

c) ⇒ a): Finally assume MP. Since $C \cap C(x_1, \ldots, x_n) = \emptyset$, we obtain that $-d(x, c) = 0$ for all $x \in C(x_1, \ldots, x_n)$ and $c \in C$. Now MP implies that indeed $d(x, c) > 0$ for all $x \in C(x_1, \ldots, x_n)$ and $c \in C$. Apply Corollary 2. \quad \square

Remark 1. We remark that Proposition 2 implies that SEP and FTAP are true in recursive constructive mathematics. For a discussion about the relation of recursive constructive mathematics to constructive mathematics we refer to [4].

Note that we obtain the following constructive version of the fundamental theorem of asset pricing.

Corollary 3. Fix a $\mathbb{R}^{m \times n}$-matrix $A$ with linearly independent rows such that $d(\xi \cdot A, x) > 0$ for all $\xi \in \mathbb{R}^m$ and $x \in X_n$. Then there exists $p \in P_n$ with $A \cdot p = 0$.

Acknowledgment The first author thanks the Japan Advanced Institute of Science and Technology for hosting him in February and March 2015.
References


Mathematisches Institut
Ludwig-Maximilians-Universität München
Theresienstraβe 39
80333 München
jberger@math.lmu.de
svindla@math.lmu.de