

Duality theory in linear optimization and its extensions — formally verified

Martin Dvorak, Vladimir Kolmogorov

2024

Abstract: Farkas established that a system of linear inequalities has a solution if and only if we cannot obtain a contradiction by taking a linear combination of the inequalities. We state and formally prove several Farkas-like theorems in Lean 4. Furthermore, we consider a linearly ordered field extended with two special elements denoted by \perp and \top where \perp is below every element and \top is above every element. We define $\perp + a = \perp = a + \perp$ for all a and we define $\top + b = \top = b + \top$ for all $b \neq \perp$. Instead of multiplication, we define scalar action $c \bullet \perp = \perp$ for every $c \geq 0$ but we define $d \bullet \top = \top$ only for $d > 0$ because $0 \bullet \top = 0$. We extend certain Farkas-like theorems to a setting where coefficients are from an extended linearly ordered field.

1 Introduction

We do not use set theory. We write everything in type theory. Juxtaposition never denotes multiplication. Juxtaposition is always function application. We often talk about linearly ordered fields. The two main examples are rationals and reals. The paper starts with famous corollaries and then proceeds to state more general theorems. Proofs are postponed to respective sections. Theorems come with names instead of numbers.

There are two famous theorems characterizing that a system of linear inequalities has a solution if and only if we cannot obtain a contradiction by taking a linear combination of the inequalities. TODO cite Farkas.

Theorem (equalityFarkas): Let I and J be finite types. Let F be a linearly ordered field. Let A be a matrix of type $(I \times J) \rightarrow F$. Let b be a vector of type $I \rightarrow F$. Exactly one of the following exists:

- nonnegative vector $x : J \rightarrow F$ such that $A \cdot x = b$
- vector $y : I \rightarrow F$ such that $A^T \cdot y \geq 0$ and $b \cdot y < 0$

```
theorem equalityFarkas (A : Matrix I J F) (b : I → F) :  
  (∃ x : J → F, 0 ≤ x ∧ A *v x = b) ≠ (∃ y : I → F, 0 ≤ AT *v y ∧ b ·v y < 0)
```

TODO cite Minkowski?

Theorem (inequalityFarkas): Let I and J be finite types. Let F be a linearly ordered field. Let A be a matrix of type $(I \times J) \rightarrow F$. Let b be a vector of type $I \rightarrow F$. Exactly one of the following exists:

- nonnegative vector $x : J \rightarrow F$ such that $A \cdot x \leq b$
- nonnegative vector $y : I \rightarrow F$ such that $A^T \cdot y \geq 0$ and $b \cdot y < 0$

```
theorem inequalityFarkas [DecidableEq I] (A : Matrix I J F) (b : I → F) :  
  (∃ x : J → F, 0 ≤ x ∧ A *v x ≤ b) ≠ (∃ y : I → F, 0 ≤ y ∧ 0 ≤ AT *v y ∧ b ·v y < 0)
```

TODO cite John von Neumann? (any maybe Oskar Morgenstern and maybe George Dantzig and maybe Albert W. Tucker?)

Theorem (StandardLP.strongDuality): Let I and J be finite types. Let F be a linearly ordered field. Let A be a matrix of type $(I \times J) \rightarrow F$. Let b be a vector of type $I \rightarrow F$. Let c be a vector of type $J \rightarrow F$. Then

$$\min \{ c \cdot x \mid x \geq 0 \wedge A \cdot x \leq b \} = - \min \{ b \cdot y \mid y \geq 0 \wedge A^T \cdot y \leq c \}$$

holds if at least one of the systems has a solution (very roughly paraphrased).

```
theorem StandardLP.strongDuality {P : StandardLP I J R} (hP : P.IsFeasible ∨ P.dualize.IsFeasible) :  
  OppositesOpt P.optimum P.dualize.optimum
```

1.1 Generalizations

The next theorem generalizes equalityFarkas to structures where multiplication does not have to be commutative. Furthermore, it supports infinitely many equations.

Theorem (coordinateFarkas): Let I be any type. Let J be a finite type. Let R be a linearly ordered division ring. Let A be an R -linear map from from $(I \rightarrow R)$ to $(J \rightarrow R)$. Let b be an R -linear map from from $(I \rightarrow R)$ to R . Exactly one of the following exists:

- nonnegative vector $x : J \rightarrow R$ such that, for all $w : I \rightarrow R$, we have $\sum_{j:J} (A w)_j \bullet x_j = b w$
- vector $y : I \rightarrow R$ such that $A y \geq 0$ and $b y < 0$

```
theorem coordinateFarkas {I J : Type*} [Fintype J] [LinearOrderedDivisionRing R]
  (A : (I → R) →l [R] J → R) (b : (I → R) →l [R] R) :
  (∃ x : J → R, 0 ≤ x ∧ ∀ w : I → R, ∑ j : J, A w j • x j = b w) ≠ (∃ y : I → R, 0 ≤ A y ∧ b y < 0)
```

In the next generalization, we replace the partially ordered module $I \rightarrow R$ by a general R -module W .

Theorem (scalarFarkas): Let J be a finite type. Let R be a linearly ordered division ring. Let W be an R -module. Let A be an R -linear map from from W to $(J \rightarrow R)$. Let b be an R -linear map from from W to R . Exactly one of the following exists:

- nonnegative vector $x : J \rightarrow R$ such that, for all $w : W$, we have $\sum_{j:J} (A w)_j \bullet x_j = b w$
- vector $y : W$ such that $A y \geq 0$ and $b y < 0$

```
theorem scalarFarkas {J : Type*} [Fintype J] [LinearOrderedDivisionRing R] [AddCommGroup W] [Module R W]
  (A : W →l [R] J → R) (b : W →l [R] R) :
  (∃ x : J → R, 0 ≤ x ∧ ∀ w : W, ∑ j : J, A w j • x j = b w) ≠ (∃ y : W, 0 ≤ A y ∧ b y < 0)
```

In the most general theorem, stated below, we replace certain occurrences of R by a linearly ordered R -module V whose order respects order on R . This result origins from TODO cite Bartl.

Theorem (fintypeFarkasBartl): Let J be a finite type. Let R be a linearly ordered division ring. Let W be an R -module. Let V be a linearly ordered R -module. Let A be an R -linear map from W to $(J \rightarrow R)$. Let b be an R -linear map from W to V . Exactly one of the following exists:

- nonnegative vector family $x : J \rightarrow V$ such that, for all $w : W$, we have $\sum_{j:J} (A w)_j \bullet x_j = b w$
- vector $y : W$ such that $A y \geq 0$ and $b y < 0$

```
theorem fintypeFarkasBartl {J : Type*} [Fintype J] [LinearOrderedDivisionRing R]
  [LinearOrderedAddCommGroup V] [Module R V] [PosSMulMono R V] [AddCommGroup W] [Module R W]
  (A : W →l [R] J → R) (b : W →l [R] V) :
  (∃ x : J → V, 0 ≤ x ∧ ∀ w : W, ∑ j : J, A w j • x j = b w) ≠ (∃ y : W, 0 ≤ A y ∧ b y < 0)
```

In the last branch, $A y \geq 0$ uses the partial order on $(J \rightarrow R)$ whereäs $b y < 0$ uses the linear order on V . Note that `fintypeFarkasBartl` subsumes `scalarFarkas` (as well as the other versions based on equality), since R can be viewed as a linearly ordered module over itself. We prove `fintypeFarkasBartl` in Section 3, which is where the heavy lifting comes.

1.2 Extensions

Until now, we have talked about known results. What follows is a new extension of the theory.

Definition: Let F be a linearly ordered field. We define an **extended** linearly ordered field F_∞ as $F \cup \{\perp, \top\}$ with the following properties. Let p and q be numbers from F . We have $\perp < p < \top$. We define addition, scalar action, and negation on F_∞ as follows:

+	\perp	q	\top
\perp	\perp	\perp	\perp
p	\perp	$p+q$	\top
\top	\perp	\top	\top

\bullet	\perp	q	\top
0	\perp	0	0
$p > 0$	\perp	$p \cdot q$	\top

-	\perp	q	\top
\perp	\perp	$-q$	\perp

When we talk about elements of F_∞ , we say that values from F are **finite**.

Informally speaking, \top represents the positive infinity, \perp represents the negative infinity, and we say that \perp is “stronger” than \top in all arithmetic operations. The surprising parts are $\perp + \top = \perp$ and $0 \bullet \perp = \perp$. Because of them, F_∞ is not a field. In fact, F_∞ is not even a group. However, F_∞ is still a densely linearly ordered abelian monoid with characteristic zero.

Theorem (extendedFarkas): Let I and J be finite types. Let F be a linearly ordered field. Let A be a matrix of type $(I \times J) \rightarrow F_\infty$. Let b be a vector of type $I \rightarrow F_\infty$. Assume that A does not have \perp and \top in the same row. Assume that A does not have \perp and \top in the same column. Assume that A does not have \top in any row where b has \top . Assume that A does not have \perp in any row where b has \perp . Exactly one of the following exists:

- nonnegative vector $x : J \rightarrow F$ such that $A \cdot x \leq b$
- nonnegative vector $y : I \rightarrow F$ such that $(-A^T) \cdot y \leq 0$ and $b \cdot y < 0$

```

theorem extendedFarkas (A : Matrix I J F $\infty$ ) (b : I  $\rightarrow$  F $\infty$ )
  (hA :  $\neg \exists i : I, (\exists j : J, A\ i\ j = \perp) \wedge (\exists j : J, A\ i\ j = \top)$ )
  (hAT :  $\neg \exists j : J, (\exists i : I, A\ i\ j = \perp) \wedge (\exists i : I, A\ i\ j = \top)$ )
  (hAb :  $\neg \exists i : I, (\exists j : J, A\ i\ j = \top) \wedge b\ i = \top$ )
  (hAb' :  $\neg \exists i : I, (\exists j : J, A\ i\ j = \perp) \wedge b\ i = \perp$ ) :
  ( $\exists x : J \rightarrow F \geq 0, A\ m^* x \leq b$ )  $\neq$  ( $\exists y : I \rightarrow F \geq 0, -A^T\ m^* y \leq 0 \wedge b\ v \cdot y < 0$ )

```

Next we define an extended notion of linear program, i.e., linear programming over extended linearly ordered fields. The implicit intention is that the linear program is to be minimized.

Definition: Let I and J be finite types. Let F be a linearly ordered field. Let A be a matrix of type $(I \times J) \rightarrow F_\infty$, let b be a vector of type $I \rightarrow F_\infty$, and c be a vector of type $J \rightarrow F_\infty$ such that the following six conditions hold:

- A does not have \perp and \top in the same row
- A does not have \perp and \top in the same column
- b does not contain \perp
- c does not contain \perp
- A does not have \top in any row where b has \top
- A does not have \perp in any column where c has \top

We say that $P = (A, b, c)$ is a **linear program** over F_∞ whose constraints are indexed by I and variables are indexed by J . We say that a nonnegative vector $x : J \rightarrow R$ is a **solution** to P if and only if $A \cdot x \leq b$. We say that P **reaches** an objective value r if and only if there exists x such that x is a solution to P and $c \cdot x = r$. We say that P is **feasible** if and only if P reaches a finite¹ value. We say that P is **bounded by** a finite value r if and only if, for every value p reached by P , we have $r \leq p$. We say that P is **unbounded** if and only if there is no finite value r such that P is bounded by r . We say that the linear program $(-A^T, c, b)$ is the **dual** of P .

Theorem (ExtendedLP.weakDuality): Let F be a linearly ordered field. Let P be a linear program over F_∞ . If P reaches p and the dual of P reaches q , then $p + q \geq 0$.

```

theorem ExtendedLP.weakDuality {P : ExtendedLP I J F}
  {p : F $\infty$ } (hP : P.Reaches p) {q : F $\infty$ } (hQ : P.dualize.Reaches q) :
  0  $\leq$  p + q

```

Definition: Let F be a linearly ordered field. Let P be a linear program over F_∞ . We define the **optimum** of P as follows. If P is feasible and unbounded, its optimum is \perp . If P is not feasible, its optimum is \top . In all other cases, we ask whether P reaches a finite value r such that P is bounded by r . If so, its optimum is r . Otherwise, P does not have optimum.²

Theorem (ExtendedLP.strongDuality):³ Let F be a linearly ordered field. Let P be a linear program over F_∞ . If P or its dual is feasible (at least one of them), then there exists p in F_∞ such that P has optimum p and the dual of P has optimum $-p$.

```

theorem ExtendedLP.strongDuality {P : ExtendedLP I J F} (hP : P.IsFeasible  $\vee$  P.dualize.IsFeasible) :
  OppositesOpt P.optimum P.dualize.optimum

```

1.3 Structure of this paper

In Section 2, we explain all underlying definitions and comment on the formalization process. In Section 3, we prove `fintypeFarkasBartl` (stated in Section 1.1), from which we obtain `equalityFarkas` as a corollary. In Section 4, we prove `extendedFarkas` (stated in Section 1.2). In Section 5, we prove `ExtendedLP.strongDuality` (stated in Section 1.2), from which we obtain `StandardLP.strongDuality` as a corollary.

Repository <https://github.com/madvorak/duality> contains full version of all definitions, statements, and proofs. They are written in a formal language called Lean 4, which provides a guarantee that every step of every proof follows from valid logical axioms.⁴ This paper attempts to be an accurate description of the **duality** project. However, in case of any discrepancy, the code shall prevail.

¹It would be perhaps more natural to say that P reaches a value different from \top . However, since \perp cannot be reached because of the way linear programming is defined, it is equivalent to our definition by reaching a finite value.

²By the end of the paper, we will have proved that optimum always exists, i.e., it cannot happen that the set of objective values reached by P has a finite infimum that is not attained. However, because we do not have the theorem now, the optimum is defined as a partial function from linear programs to F_∞ .

³For simplicity, we rephrased the theorem without mentioning partial functions. Also note that it would be incorrect to say the following: P has optimum p , the dual of P has optimum q , and $p + q = 0$. It would fail for unbounded linear programs because the arithmetics of F_∞ defines $\top + \perp = \perp$.

⁴The only axioms used in our proofs are `propext`, `Classical.choice`, and `Quot.sound`, which you can check by the `#print axioms` command.

2 Preliminaries

2.1 We start with a review of algebraic typeclasses that our project depends on

Additive semigroup is a structure on any type with addition where the addition is associative:

```
class AddSemigroup (G : Type u) extends Add G where
  add_assoc : ∀ a b c : G, (a + b) + c = a + (b + c)
```

Additive monoid is an additive semigroup with the zero element, thanks to which we can define a scalar multiplication by the natural numbers (TODO AddZeroClass):

```
class AddMonoid (M : Type u) extends AddSemigroup M, AddZeroClass M where
  nsmul : ℕ → M → M
  nsmul_zero : ∀ x : M, nsmul 0 x = 0
  nsmul_succ : ∀ (n : ℕ) (x : M), nsmul (n + 1) x = nsmul n x + x
```

Subtractive monoid is an additive monoid that adds two more operations (unary and binary minus) that satisfy some basic properties:

```
class SubNegMonoid (G : Type u) extends AddMonoid G, Neg G, Sub G where
  sub := SubNegMonoid.sub'
  sub_eq_add_neg : ∀ a b : G, a - b = a + -b
  zsmul : ℤ → G → G
  zsmul_zero' : ∀ a : G, zsmul 0 a = 0
  zsmul_succ' (n : ℕ) (a : G) : zsmul (Int.ofNat n.succ) a = zsmul (Int.ofNat n) a + a
  zsmul_neg' (n : ℕ) (a : G) : zsmul (Int.negSucc n) a = -(zsmul n.succ a)
```

Additive group is a subtractive monoid in which the unary minus acts as an inverse with respect to addition:

```
class AddGroup (A : Type u) extends SubNegMonoid A where
  add_left_neg : ∀ a : A, -a + a = 0
```

Abelian group is defined as an additive group that is a commutative additive monoid at the same time (TODO AddCommMonoid):

```
class AddCommGroup (G : Type u) extends AddGroup G, AddCommMonoid G
```

Ring is defined as a semiring that is an abelian group at the same time and has 1 that behaves well (TODO Semiring):

```
class Ring (R : Type u) extends Semiring R, AddCommGroup R, AddGroupWithOne R
```

Division ring is a ring with a lot of extra requirements (TODOs DivInvMonoid, Nontrivial, NNRatCast, RatCast):

```
class DivisionRing (α : Type*) extends Ring α, DivInvMonoid α, Nontrivial α, NNRatCast α, RatCast α where
  mul_inv_cancel : ∀ (a : α), a ≠ 0 → a * a-1 = 1
  inv_zero : (0 : α)-1 = 0
  nnratCast := NNRat.castRec
```

We define a linearly ordered division ring as a division ring that is a linearly ordered ring at the same time (TODO all about order):

```
class LinearOrderedDivisionRing (R : Type*) extends LinearOrderedRing R, DivisionRing R
```

Linearly ordered field is defined as a linearly ordered commutative ring that is a field at the same time (TODO Field):

```
class LinearOrderedField (α : Type*) extends LinearOrderedCommRing α, Field α
```

Note that LinearOrderedDivisionRing is not a part of the algebraic hierarchy provided by Mathlib, hence LinearOrderedField does not inherit LinearOrderedDivisionRing, thus we provide a custom instance that converts LinearOrderedField to LinearOrderedDivisionRing:

```
instance LinearOrderedField.toLinearOrderedDivisionRing {F : Type*} [instF : LinearOrderedField F] :
  LinearOrderedDivisionRing F := { instF with }
```

This instance is needed for the step from coordinateFarkas to equalityFarkas.

2.2 Extended linearly ordered fields

Given any type F , we construct $F \cup \{\perp, \top\}$ as follows:

```
def Extend (F : Type*) := WithBot (WithTop F)
```

From now on we assume that F is a linearly ordered field:

```
variable {F : Type*} [LinearOrderedField F]
```

The following instance defines how addition and comparison behaves on F_∞ and automatically generates a proof that F_∞ forms a linearly ordered abelian monoid:

```
instance : LinearOrderedAddCommMonoid (Extend F) :=
  inferInstanceAs (LinearOrderedAddCommMonoid (WithBot (WithTop F)))
```

The following definition embeds F in F_∞ and registers this canonical embedding as a type coercion:

```
@[coe] def toE : F → (Extend F) := some ∘ some
instance : Coe F (Extend F) := ⟨toE⟩
```

Unary minus on F_∞ is defined as follows:

```
def neg : Extend F → Extend F
| ⊥ => ⊤
| ⊤ => ⊥
| (x : F) => toE (-x)
instance : Neg (Extend F) := ⟨EF.neg⟩
```

In the file `FarkasSpecial.lean` and everything downstream, we have the notation

F_∞

for F_∞ and also the notation

$F_{\geq 0}$

for all nonnegative elements of F . TODO define scalar action and explain `SMulZeroClass`.

2.3 Vectors and matrices

We distinguish two types of vectors; implicit vectors and explicit vectors. Implicit vectors are members of a vector space; they don't have any internal structure. Explicit vectors are functions from coordinates to values. The set of coordinates needn't be ordered. Matrices live next to explicit vectors. They are also functions; they take a row index and a column index and they output a value at the given spot. Neither the row indices nor the column vertices are required to form an ordered set. That's why multiplication between matrices and vectors is defined only in structures where addition forms a commutative semigroup. Consider the following example:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} = \begin{pmatrix} ? \\ - \end{pmatrix}$$

We don't know whether the value at the question mark is equal to $(1 \cdot 7 + 2 \cdot 8) + 3 \cdot 9$ or to $(2 \cdot 8 + 1 \cdot 7) + 3 \cdot 9$ or to any other ordering of summands. This is why commutativity of addition is necessary for the definition to be valid. On the other hand, we don't assume any property of multiplication in the definition of multiplication between matrices and vectors; they don't even have to be of the same type; we only require the elements of the vector to have an action on the elements of the matrix (this is not a typo — normally, we would want matrices to have an action on vectors — not in our work).

2.3.1 Mathlib definitions versus our definitions

Mathlib defines dot product (i.e., a product of an explicit vector with an explicit vector) as follows:

```
def Matrix.dotProduct [Fintype m] [Mul α] [AddCommMonoid α] (v w : m → α) : α :=
  ∑ i, v i * w i
```

Mathlib defines a product of a matrix with an explicit vector as follows:

```
def Matrix.mulVec [Fintype n] [NonUnitalNonAssocSemiring α] (M : Matrix m n α) (v : n → α) : m → α
| i => (fun j => M i j) ·_v v
```

Mildly confusingly, `m` and `n` are type variables here, not natural numbers. Note that, when multiplying two vectors, the left vector is not transposed — a vector is not defined as a special case of matrix, and transposition is not defined as an operation on explicit vectors, only on matrices.

Infix notation for these two operations is defined as follows (the keyword `infixl` when defining an operator \circ means that $x \circ y \circ z$ gets parsed as $(x \circ y) \circ z$ whether or not it makes sense, whereäs the keyword `infixr` when defining an operator \circ means that $A \circ B \circ v$ gets parsed as $A \circ (B \circ v)$ whether or not it makes sense):

```
scoped infixl:72 " ·_v " => Matrix.dotProduct
scoped infixr:73 " *_v " => Matrix.mulVec
```

These definitions are sufficient for stating results based on linearly ordered fields. However, our results from Section 1.3 require a more general notion of multiplication between matrices and vectors. They are defined in the section `hetero_matrix_products_defs` as follows:

```
variable {α γ : Type*} [AddCommMonoid α] [SMul γ α]
```

```
def Matrix.dotProd (v : I → α) (w : I → γ) : α :=
  ∑ i : I, w i • v i
infixl:72 " ·_v " => Matrix.dotProd
```

```
def Matrix.mulWeig (M : Matrix I J α) (w : J → γ) (i : I) : α :=
  M i ·_v w
infixr:73 " *_m " => Matrix.mulWeig
```

We start by declaring that α and γ are types from any universe (not necessarily both from the same universe). We require that α forms an abelian monoid and that γ has a scalar action on α . In this setting, we can instantiate α with F_∞ and γ with F for any linearly ordered field F .

For explicit vectors $v : I \rightarrow \alpha$ and $w : I \rightarrow \gamma$, we define their product of type α as follows. Every element of v gets multiplied from left by an element of w on the same index. Then we sum them all together (in unspecified order). For a matrix $M : (I \times J) \rightarrow \alpha$ and a vector $w : J \rightarrow \gamma$, we define their product of type $I \rightarrow \alpha$ as a function that takes an index i and outputs the dot product between the i -th row of M and the vector w .

Beware that the arguments (both in the function definition and in the infix notation) come in the opposite order from how scalar action is written. We recommend a mnemonic “vector times weights” for $v \cdot w$ and “matrix times weights” for $M \cdot w$ where arguments come in alphabetical order.

In the infix notation, you can distinguish between the standard Mathlib definitions and our definitions by observing that Mathlib operators put the letter v to the right of the symbol whereas our operators put a letter to the left of the symbol.

We find it unfortunate that the Mathlib name of `dotProduct` is prefixed by the `Matrix` namespace. While `Matrix.mulVec` allows for the use of dot notation in places where infix notation is not used, the full name `Matrix.dotProduct` only clutters the code. Even though we don’t like it, we decided to namespace our definitions in the same way for consistency.

Since we have new definitions, we have to rebuild all API (a lot of lemmas) for `Matrix.dotProd` and `Matrix.mulWeig` from scratch. This process was very tiresome. We decided not to develop a full reusable library, but prove only those lemmas we wanted to use in our project. For similar reasons, we did not generalize the Mathlib definition of “vector times matrix”, as “matrix times vector” was all we needed. It was still a lot of lemmas.

2.4 Linear programming

Extended linear programs are defined as follows:

```
structure ExtendedLP (I J F : Type*) [LinearOrderedField F] where
  /-- The left-hand-side matrix. -/
  A : Matrix I J F∞
  /-- The right-hand-side vector. -/
  b : I → F∞
  /-- The objective function coefficients. -/
  c : J → F∞
  /-- No ‘⊥’ and ‘⊤’ in the same row. -/
  hAi : ¬∃ i : I, (∃ j : J, A i j = ⊥) ∧ (∃ j : J, A i j = ⊤)
  /-- No ‘⊥’ and ‘⊤’ in the same column. -/
  hAj : ¬∃ j : J, (∃ i : I, A i j = ⊥) ∧ (∃ i : I, A i j = ⊤)
  /-- No ‘⊥’ in the right-hand-side vector. -/
  hbi : ¬∃ i : I, b i = ⊥
  /-- No ‘⊥’ in the objective function coefficients. -/
  hcj : ¬∃ j : J, c j = ⊥
  /-- No ‘⊤’ in the row where the right-hand-side vector has ‘⊤’. -/
  hAb : ¬∃ i : I, (∃ j : J, A i j = ⊤) ∧ b i = ⊤
  /-- No ‘⊥’ in the column where the objective function has ‘⊤’. -/
  hAc : ¬∃ j : J, (∃ i : I, A i j = ⊥) ∧ c j = ⊤
```

Solution is defined as follows:

```
def ExtendedLP.IsSolution (P : ExtendedLP I J F) (x : J → F≥0) : Prop :=
  P.A * x ≤ P.b
```

Reaching a value is defined as follows:

```
def ExtendedLP.Reaches (P : ExtendedLP I J F) (r : F∞) : Prop :=
  ∃ x : J → F≥0, P.IsSolution x ∧ P.c v · x = r
```

Feasibility is defined as follows:

```
def ExtendedLP.IsFeasible (P : ExtendedLP I J F) : Prop :=
  ∃ p : F, P.Reaches (toE p)
```

Being bounded by a value (from below – we always minimize) is defined as follows:

```
def ExtendedLP.IsBoundedBy (P : ExtendedLP I J F) (r : F) : Prop :=
  ∀ p : F∞, P.Reaches p → r ≤ p
```

Being unbounded is defined as follows:

```
def ExtendedLP.IsUnbounded (P : ExtendedLP I J F) : Prop :=
  ¬∃ r : F, P.IsBoundedBy r
```

The following definition says how linear programs are dualized:

```
def ExtendedLP.dualize (P : ExtendedLP I J F) : ExtendedLP J I F :=
  ⟨-P.A⊤, P.c, P.b, by aeplly P.hAj, by aeplly P.hAi, P.hcj, P.hbi, by aeplly P.hAc, by aeplly P.hAb⟩
```

The definition of optimum is, sadly, very complicated:

```

noncomputable def ExtendedLP.optimum (P : ExtendedLP I J F) : Option F∞ :=
  if P.IsFeasible then
    if P.IsUnbounded then
      some ⊥ -- unbounded means that the minimum is '⊥'
    else
      if hf : ∃ r : F, P.Reaches (toE r) ∧ P.IsBoundedBy r then
        some (toE hf.choose) -- the minimum is finite
      else
        none -- invalid finite value (infimum is not attained)
    else
      some ⊤ -- infeasible means that the minimum is '⊤'

```

Finally, we define what opposite values are:

```

def OppositesOpt : Option F∞ → Option F∞ → Prop
| (p : F∞), (q : F∞) => p = -q -- opposite values; includes '⊥ = -⊤' and '⊤ = -⊥'
| _ , _ => False -- namely 'OppositesOpt none none = False'

```

3 Proving the Farkas-Bartl theorem

We prove `finFarkasBartl` and, in the end, we obtain `fintypeFarkasBartl` as corollary.

Theorem (finFarkasBartl): Let n be a natural number. Let R be a linearly ordered division ring. Let W be an R -module. Let V be a linearly ordered R -module. Let A be an R -linear map from W to $([n] \rightarrow R)$. Let b be an R -linear map from W to V . Exactly one of the following exists:

- nonnegative vector family $x : [n] \rightarrow V$ such that, for all $w : W$, we have $\sum_{j:[n]} (A w)_j \bullet x_j = b w$
- vector $y : W$ such that $A y \geq 0$ and $b y < 0$

The only difference is that `finFarkasBartl` uses $[n] = \{0, \dots, n-1\}$ instead of an arbitrary (unordered) finite type J .

Proof idea: We first prove that both cannot exist at the same time. Assume we have x and y of said properties. We plug y for w and obtain $\sum_{j:[n]} (A y)_j \bullet x_j = b y$. On the left-hand side, we have a sum of nonnegative vectors, which contradicts $b y < 0$.

We prove “at least one exists” by induction on n . If $n = 0$ then $A y \geq 0$ is a tautology. We consider b . Either b maps everything to the zero vector, which allows x to be the empty vector family, or some w gets mapped to a nonzero vector, where we choose y to be either w or $(-w)$. Since V is linearly ordered, one of them satisfies $b y < 0$. Now we precisely state how the induction step will be.

Lemma (industepFarkasBartl): Let m be a natural number. Let R be a linearly ordered division ring. Let W be an R -module. Let V be a linearly ordered R -module. Assume (induction hypothesis) that for all R -linear maps $A_0 : W \rightarrow ([m] \rightarrow R)$ and $b_0 : W \rightarrow V$, the formula “ $\forall y_0 : W, A_0 y_0 \geq 0 \implies b_0 y_0 \geq 0$ ” implies existence of a nonnegative vector family $x_0 : [m] \rightarrow V$ such that, for all $w_0 : W$, $\sum_{i:[m]} (A_0 w_0)_i \bullet (x_0)_i = b_0 w_0$. Let A be an R -linear map from W to $([m+1] \rightarrow R)$. Let b be an R -linear map from W to V . Assume that, for all $y : W$, $A y \geq 0$ implies $b y \geq 0$. We claim there exists a nonnegative vector family $x : [m+1] \rightarrow V$ such that, for all $w : W$, we have $\sum_{i:[m+1]} (A w)_i \bullet x_i = b w$. TODO names like A_0 don’t work well with the “subscript notation” on paper.

Proof idea: Let $A_{<m}$ roughly mean $A \upharpoonright [m]$. To be more precise, $A_{<m}$ is a function that maps $(w : W)$ to $(A w) \upharpoonright [m]$, i.e., $A_{<m}$ is an R -linear map from W to $([m] \rightarrow R)$ that behaves exactly like A where it is defined. We distinguish two cases. If, for all $y : W$, the inequality $A_{<m} y \geq 0$ implies $b y \geq 0$, then plug $A_{<m}$ for A_0 , obtain x_0 , and construct a vector family x such that $x_m = 0$ and otherwise x copies x_0 . We easily check that x is nonnegative and that $\sum_{i:[m+1]} (A w)_i \bullet x_i = b w$ holds.

In the second case, we have y' such that $A_{<m} y' \geq 0$ holds but $b y' < 0$ also holds. We realize that $(A y')_m < 0$. We now declare $y := (A y')_m \bullet y'$ and observe $(A y)_m = 1$. We establish the following facts (proofs are omitted):

- for all $w : W$, we have $A (w - ((A w)_m \bullet y)) = 0$
- for all $w : W$, the inequality $A_{<m} (w - ((A w)_m \bullet y)) \geq 0$ implies $b (w - ((A w)_m \bullet y)) \geq 0$
- for all $w : W$, the inequality $A_{<m} w - A_{<m} ((A w)_m \bullet y) \geq 0$ implies $b w - b ((A w)_m \bullet y) \geq 0$
- for all $w : W$, the inequality $(A_{<m} - (z \mapsto (A z)_m \bullet (A_{<m} y))) w \geq 0$ implies $(b - (z \mapsto (A z)_m \bullet (b y))) w \geq 0$

We observe that $A_0 := A_{<m} - (z \mapsto (A z)_m \bullet (A_{<m} y))$ and $b_0 := b - (z \mapsto (A z)_m \bullet (b y))$ are R -linear maps. Thanks to the last fact, we can apply induction hypothesis to A_0 and b_0 . We obtain a nonnegative vector family x' such that, for all $w_0 : W$, $\sum_{i:[m]} (A_0 w_0)_i \bullet x'_i = b_0 w_0$. It remains to construct a nonnegative vector family $x : [m+1] \rightarrow V$ such that, for all $w : W$, we have $\sum_{i:[m+1]} (A w)_i \bullet x_i = b w$. We choose $x_m = b y - \sum_{i:[m]} (A_{<m} y)_i \bullet x'_i$ and otherwise x copies x' . We check that our x has the required properties. Qed.

We complete the proof of `finFarkasBartl` by applying `industepFarkasBartl` to $A_{\leq n}$ and b . Finally, we obtain `fintypeFarkasBartl` from `finFarkasBartl` using some boring mechanisms regarding equivalence between finite types.

4 Extended Farkas theorem

Theorem (inequalityFarkas_neg): Let I and J be finite types. Let F be a linearly ordered field. Let A be a matrix of type $(I \times J) \rightarrow F$. Let b be a vector of type $I \rightarrow F$. Exactly one of the following exists:

- nonnegative vector $x : J \rightarrow F$ such that $A \cdot x \leq b$
- nonnegative vector $y : I \rightarrow F$ such that $(-A^T) \cdot y \leq 0$ and $b \cdot y < 0$

Obviously, inequalityFarkas_neg is an immediate corollary of inequalityFarkas.

Theorem (extendedFarkas): Let I and J be finite types. Let F be a linearly ordered field. Let A be a matrix of type $(I \times J) \rightarrow F_\infty$. Let b be a vector of type $I \rightarrow F_\infty$. Assume that A does not have \perp and \top in the same row. Assume that A does not have \perp and \top in the same column. Assume that A does not have \top in any row where b has \top . Assume that A does not have \perp in any row where b has \perp . Exactly one of the following exists:

- nonnegative vector $x : J \rightarrow F$ such that $A \cdot x \leq b$
- nonnegative vector $y : I \rightarrow F$ such that $(-A^T) \cdot y \leq 0$ and $b \cdot y < 0$

(restated)

4.1 Proof idea

We need to do the following steps in given order:

1. Delete all rows of both A and b where A has \perp or b has \top (they are tautologies).
2. Delete all columns of A that contain \top (they force respective variables to be zero).
3. If b contains \perp , then $A \cdot x \leq b$ cannot be satisfied, but $y = 0$ satisfies $(-A^T) \cdot y \leq 0$ and $b \cdot y < 0$. Stop here.
4. Assume there is no \perp in b . Use inequalityFarkas_neg. In either case, extend x or y with zeros on all deleted positions.

4.2 Counterexamples

If A has \perp and \top in the same row, it may happen that both x and y exist:

$$A = \begin{pmatrix} \perp & \top \\ 0 & -1 \end{pmatrix} \quad b = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \quad x = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

If A has \perp and \top in the same column, it may happen that both x and y exist:

$$A = \begin{pmatrix} \perp \\ \top \end{pmatrix} \quad b = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad x = \begin{pmatrix} 0 \end{pmatrix} \quad y = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

If A has \top in a row where b has \top , it may happen that both x and y exist:

$$A = \begin{pmatrix} \top \\ -1 \end{pmatrix} \quad b = \begin{pmatrix} \top \\ -1 \end{pmatrix} \quad x = \begin{pmatrix} 1 \end{pmatrix} \quad y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

If A has \perp in a row where b has \perp , it may happen that both x and y exist:

$$A = \begin{pmatrix} \perp \end{pmatrix} \quad b = \begin{pmatrix} \perp \end{pmatrix} \quad x = \begin{pmatrix} 1 \end{pmatrix} \quad y = \begin{pmatrix} 0 \end{pmatrix}$$

5 Extended strong LP duality

We start with the weak duality and then move to the strong duality. We will use extendedFarkas in several places.

Theorem (weakDuality): Let F be a linearly ordered field. Let P be a linear program over F_∞ . If P reaches p and the dual of P reaches q , then $p + q \geq 0$. (restated)

Proof idea: There is a vector x such that $A \cdot x \leq b$ and $c \cdot x = p$. Apply extendedFarkas to the following matrix and vector:

$$\begin{pmatrix} A \\ c \end{pmatrix} \quad \begin{pmatrix} b \\ p \end{pmatrix}$$

Lemma (infeasible_of_unbounded): If a linear program P is unbounded, the dual of P cannot be feasible.

Proof idea: Assume that P is unbounded, but the dual of P is feasible. Obtain contradiction using weakDuality.

Lemma (unbounded_of_reaches_le): Let F be a linearly ordered field. Let P be a linear program over F_∞ . Assume that for each s in F there exists p in F_∞ such that P reaches p and $p \leq s$. Then P is unbounded.

Proof idea: It suffices to prove that for each r in F there exists p' in F_∞ such that P reaches p' and $p' < r$. Apply the assumption to $r-1$.

Lemma (unbounded_of_feasible_of_neg): Let F be a linearly ordered field. Let P be a linear program over F_∞ that is feasible. Let x_0 be a nonnegative vector such that $c \cdot x_0 < 0$ and $A \cdot x_0 + 0 \bullet (-b) \leq 0$. Then P is unbounded.

Proof idea: There is a nonnegative vector x_p such that $A \cdot x_p \leq b$ and $c \cdot x_p = e$ for some e in F_∞ . We apply unbounded_of_reaches_le. In case $e \leq s$, we use x_p and we are done. Otherwise, consider what $c \cdot x_0$ equals to. We cannot have $c \cdot x_0 = \perp$ because c does not contain \perp . We cannot have $c \cdot x_0 = \top$ because $c \cdot x_0 < 0$. Hence $c \cdot x_0 = d$ for some d in F . Observe that the fraction $\frac{s-e}{d}$ is well defined and it is positive. Use $x_p + \frac{s-e}{d} \bullet x_0$.

Lemma (strongDuality_aux): Let P be a linear program such that P is feasible and the dual of P is also feasible. There is a value p reached by P and a value q reached by the dual of P such that $p + q \leq 0$.

Proof idea: TODO.

Lemma (strongDuality_of_both_feasible): Let P be a linear program such that P is feasible and the dual of P is also feasible. There is a finite value r such that P reaches $-r$ and the dual of P reaches r .

Proof idea: From strongDuality_aux we have a value p reached by P and a value q reached by the dual of P such that $p + q \leq 0$. We apply weakDuality to p and q to obtain $p + q \geq 0$. We set $r := q$.

Lemma (unbounded_of_feasible_of_infeasible): Let P be a linear program such that P is feasible but the dual of P is not feasible. Then P is unbounded.

Proof idea: TODO.

Lemma (optimum_unique): Let P be a linear program. Let r be a valued reached by P such that P is bounded by r . Let s be a valued reached by P such that P is bounded by s . Then $r = s$.

Proof idea: TODO.

Lemma (optimum_eq_of_reaches_bounded): Let P be a linear program. Let r be a valued reached by P such that P is bounded by r . Then the optimum of P is r .

Proof idea: Apply the axiom of choice to the definition of optimum and use optimum_unique.

Lemma (strongDuality_of_prim_feas): Let P be a linear program that is feasible. The strong duality holds.

Proof idea: TODO.

Theorem (optimum_neq_none): Every linear program has optimum.

Proof idea: If a linear program P is feasible, the existence of optimum follows from strongDuality_of_prim_feas. Otherwise, the optimum of P is \top by definition.

Lemma (dualize_dualize): Let P be a linear program. The dual of the dual of P is exactly P .

Proof idea: $-(-A^T)^T = A$

Lemma (strongDuality_of_dual_feas): Let P be a linear program whose dual is feasible. The strong duality holds.

Proof idea: Apply strongDuality_of_prim_feas to the dual of P and use dualize_dualize.

Theorem (strongDuality): Let F be a linearly ordered field. Let P be a linear program over F_∞ . If P or its dual is feasible (at least one of them), then there exists p in F_∞ such that P has optimum p and the dual of P has optimum $-p$. (restated)

Proof idea: Use strongDuality_of_prim_feas or strongDuality_of_dual_feas.

5.1 Counterexample (first two conditions)

If A has \perp and \top in the same column, the following may happen:

$$A = \begin{pmatrix} \perp \\ \top \end{pmatrix} \quad b = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad c = (0)$$

If A has \perp and \top in the same row, the following may happen:

$$A = \begin{pmatrix} \top & \perp \end{pmatrix} \quad b = (0) \quad c = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

The two linear programs above are duals of each other. Yet the optimum of the former is 0 but the optimum of the latter is \perp .

5.2 Counterexample (middle two conditions)

If b contains \perp , the following may happen:

$$A = \begin{pmatrix} \perp \end{pmatrix} \qquad b = \begin{pmatrix} \perp \end{pmatrix} \qquad c = \begin{pmatrix} 0 \end{pmatrix}$$

If c contains \perp , the following may happen:

$$A = \begin{pmatrix} \top \end{pmatrix} \qquad b = \begin{pmatrix} 0 \end{pmatrix} \qquad c = \begin{pmatrix} \perp \end{pmatrix}$$

The two linear programs above are duals of each other. Yet the optimum of the former is 0 but the optimum of the latter is \perp .

5.3 Counterexample (last two conditions)

If A has \top in a row where b has \top , the following may happen:

$$A = \begin{pmatrix} \top \\ -1 \end{pmatrix} \qquad b = \begin{pmatrix} \top \\ -1 \end{pmatrix} \qquad c = \begin{pmatrix} 0 \end{pmatrix}$$

If A has \perp in a column where c has \top , the following may happen:

$$A = \begin{pmatrix} \perp & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 0 \end{pmatrix} \qquad c = \begin{pmatrix} \top \\ -1 \end{pmatrix}$$

The two linear programs above are duals of each other. Yet the optimum of the former is 0 but the optimum of the latter is \perp .

6 Related work

TODO presents an overcomplicated proof in Isabelle by analyzing the Simplex algorithm that already had been formally verified. It took them 30 pages to get to the basic Farkas; no generalization was provided.

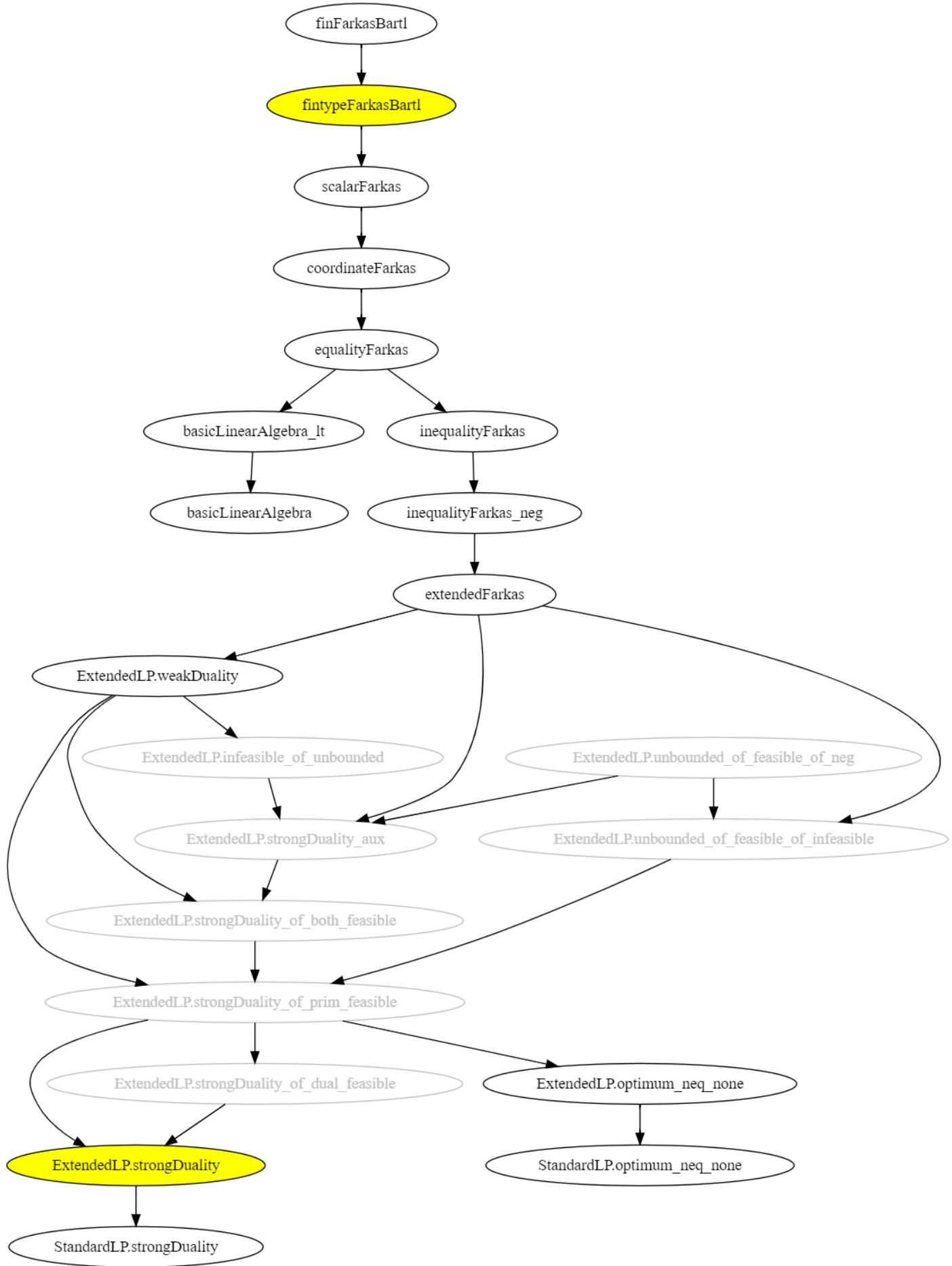
In Lean, it would be possible to prove Farkas for reals using the Hahn-Banach separation theorem. However, we do not yet know that the set of feasible solutions is closed.

TODO.

7 Conclusion

We formally verified several Farkas-like theorems in Lean 4. We extended the existing theory to a new setting where some coefficient can carry infinite values. We realized that the abstract work with modules over linearly ordered division rings and linear maps between them was fairly easy to carry on in Lean 4 thanks to the library Mathlib that is perfectly suited for such tasks. In contrast, manipulation with matrices got tiresome whenever we needed a not-fully-standard operation. It turns out Lean 4 cannot automate case analyses unless they take place in the “outer layers” of formulas. Summation over subtypes and summation of conditional expression made us developed a lot of ad-hoc machinery which we would have preferred to be handled by existing tactics. Another area where Lean 4 is not yet helpful is the search for counterexamples. Despite these difficulties, we find Lean 4 to be an extremely valuable tool for elegant expressions of mathematical formulas and for proving them formally.

8 Appendix: dependencies between theorems



Theorems are in black. Selected lemmas are in gray. What we consider to be the main theorems are denoted by yellow background.