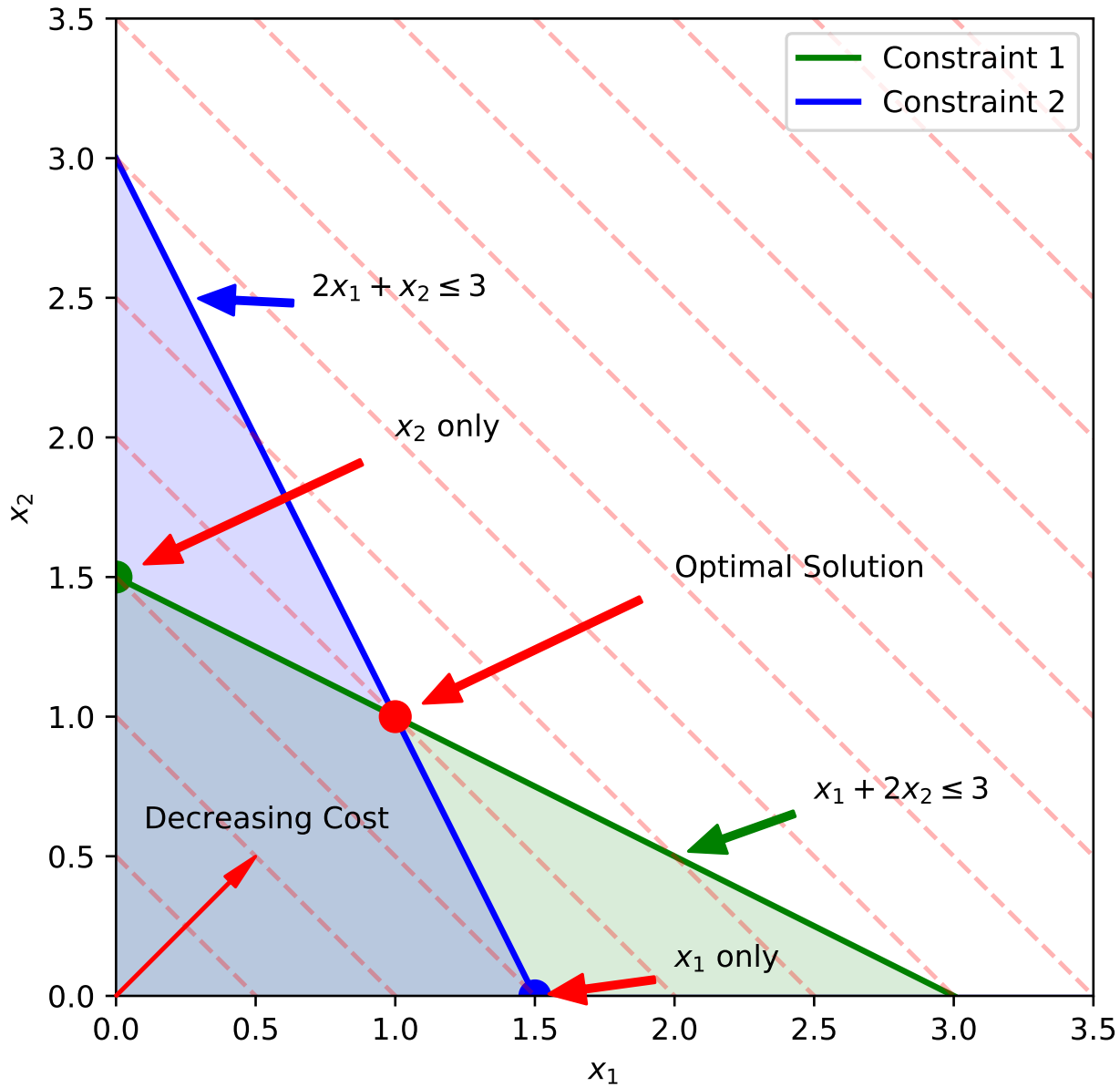
Linear Algebra Background

Stefano Gualandi
Università di Pavia, Dipartimento di Matematica

email: stefano.gualandi@unipv.it
twitter: @famo2spaghi
blog: http://stegua.github.com

Intuition is not enough...



Notation

- All vectors $\boldsymbol{x} \in \mathbb{R}^n$ are column vectors, unless explicitly stated.
- We use $\mathbf{0}_n$ to denote the vector with n zeros (or $\mathbf{0}$, dimension from context)
- The i-th unit vector is the vector with all 0 except the i-th component.
- The *inner product* of two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ is: $\boldsymbol{x}^\intercal \boldsymbol{y} = \boldsymbol{y}^\intercal \boldsymbol{x} = \sum_{i=1}^{n} x_i y_i$.
- Two vectors are **orthogonal** if $\boldsymbol{x}^\intercal \boldsymbol{y} = 0$.
- The Euclidean norm is $||\boldsymbol{x}||_2 = \sqrt{\boldsymbol{x}^\intercal \boldsymbol{x}}$.
- *Schwartz inequality*: $|\boldsymbol{x}^\intercal \boldsymbol{y}| \leq ||\boldsymbol{x}|| \cdot ||\boldsymbol{y}||$ (with equality if and only if $\boldsymbol{x} = \lambda \boldsymbol{y}$)
- If $\boldsymbol{A}$ is an $m \times n$ matrix, then $\boldsymbol{A}_j$ is its j-th column and $\boldsymbol{a}_i$ the i-th row (SCRIVI).
- If $\boldsymbol{x}$ is a vector, $\boldsymbol{x} \geq 0$ means that every component of $\boldsymbol{x}$ is non negative.
  If $\boldsymbol{A}$ is a matrix, then $\boldsymbol{A} \geq 0$ has a similar meaning.

Definition 1 (Linearly independence)

Given the vectors $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^K \in \mathbb{R}^n$, we say that they are linearly dependent if there exist real numbers $a_1, \ldots, a_K$, not all of them zero, such that $\sum_{i=1}^{K} a_k \boldsymbol{x}^k = 0$.

Otherwise, the K vectors are called linearly independent.

Definition 1 (Linearly independence)

Given the vectors $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^K \in \mathbb{R}^n$, we say that they are linearly dependent if there exist real numbers $a_1, \ldots, a_K$, not all of them zero, such that $\sum_{i=1}^{K} a_k \boldsymbol{x}^k = 0$.

Otherwise, the K vectors are called linearly independent.

Definition 2 (Matrix inversion)

Let $\boldsymbol{A}$ be a square matrix. If there exists a square matrix $\boldsymbol{B}$ of the same dimension, such that $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$, we say that $\boldsymbol{A}$ is invertible or non-singular.

The matrix $\boldsymbol{B}$ is the *inverse* of $\boldsymbol{A}$, and is denoted by $\boldsymbol{A}^{-1}$.

Theorem 3

*Let **A** be a square matrix. Then the following statements are equivalent:*

1. *The matrix **A** is invertible.*

Theorem 3

Let $\boldsymbol{A}$ be a square matrix. Then the following statements are equivalent:

1. *The matrix $\boldsymbol{A}$ is invertible.*
2. *The matrix $\boldsymbol{A}^{\mathsf{T}}$ is invertible.*

Theorem 3

Let $\boldsymbol{A}$ be a square matrix. Then the following statements are equivalent:

1. *The matrix $\boldsymbol{A}$ is invertible.*
2. *The matrix $\boldsymbol{A}^{\mathsf{T}}$ is invertible.*
3. *The determinant of $\boldsymbol{A}$ is non-zero.*

Theorem 3

Let $\boldsymbol{A}$ be a square matrix. Then the following statements are equivalent:

1. *The matrix $\boldsymbol{A}$ is invertible.*
2. *The matrix $\boldsymbol{A}^{\mathsf{T}}$ is invertible.*
3. *The determinant of $\boldsymbol{A}$ is non-zero.*
4. *The rows of $\boldsymbol{A}$ are linearly independent.*

Theorem 3

Let $\boldsymbol{A}$ be a square matrix. Then the following statements are equivalent:

1. *The matrix $\boldsymbol{A}$ is invertible.*
2. *The matrix $\boldsymbol{A}^{\mathsf{T}}$ is invertible.*
3. *The determinant of $\boldsymbol{A}$ is non-zero.*
4. *The rows of $\boldsymbol{A}$ are linearly independent.*
5. *The columns of $\boldsymbol{A}$ are linearly independent.*

Theorem 3

Let $\boldsymbol{A}$ be a square matrix. Then the following statements are equivalent:

1. *The matrix $\boldsymbol{A}$ is invertible.*
2. *The matrix $\boldsymbol{A}^{\mathsf{T}}$ is invertible.*
3. *The determinant of $\boldsymbol{A}$ is non-zero.*
4. *The rows of $\boldsymbol{A}$ are linearly independent.*
5. *The columns of $\boldsymbol{A}$ are linearly independent.*
6. *For every vector $\boldsymbol{b}$, the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique solution.*

Theorem 3

Let $\boldsymbol{A}$ be a square matrix. Then the following statements are equivalent:

1. *The matrix $\boldsymbol{A}$ is invertible.*
2. *The matrix $\boldsymbol{A}^{\mathsf{T}}$ is invertible.*
3. *The determinant of $\boldsymbol{A}$ is non-zero.*
4. *The rows of $\boldsymbol{A}$ are linearly independent.*
5. *The columns of $\boldsymbol{A}$ are linearly independent.*
6. *For every vector $\boldsymbol{b}$, the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique solution.*
7. *There exists some vector $\boldsymbol{b}$ such that the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique solution.*

Definition 4 (Subspace)

A nonempty subset S of $\mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if $(ax + by) \in S$ for every $x, y \in S$ and every $a, b \in \mathbb{R}$.

If in addition, $S \neq \mathbb{R}^n$, we say that S is a proper subspace.

Definition 4 (Subspace)

A nonempty subset S of $\mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if $(\boldsymbol{ax} + \boldsymbol{by}) \in S$ for every $\boldsymbol{x}, \boldsymbol{y} \in S$ and every $a, b \in \mathbb{R}$.

If in addition, $S \neq \mathbb{R}^n$, we say that S is a proper subspace.

Definition 5 (Span)

The span of a finite number of vectors $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^K$ is the subspace of $\mathbb{R}^n$ defined as the set of vectors $\boldsymbol{y} = \sum_{k=1}^{K} a_k \boldsymbol{x}^k$, where each $a_k \in \mathbb{R}$.

Any such $\boldsymbol{y}$ is called a linear combination of $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^K$.

Definition 4 (Subspace)

A nonempty subset S of $\mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if $(ax + by) \in S$ for every $x, y \in S$ and every $a, b \in \mathbb{R}$.

If in addition, $S \neq \mathbb{R}^n$, we say that S is a proper subspace.

Definition 5 (Span)

The span of a finite number of vectors $x^1, \ldots, x^K$ is the subspace of $\mathbb{R}^n$ defined as the set of vectors $y = \sum_{k=1}^{K} a_k x^k$, where each $a_k \in \mathbb{R}$.

Any such y is called a linear combination of $x^1, \ldots, x^K$.

Definition 6 (Basis)

Given a subspace S of $\mathbb{R}^n$, with $S \neq \{0\}$, a basis of S is a collection of vectors that are linearly independent and whose span is equal to S.

Every basis of a given subspace has the same number of vectors and this number is called the dimension of the subspace.

Definition 7 (Subspace)

A nonempty subset S of $\mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if $ax + by \in S$ for every $x, y \in S$ and every $a, b \in \mathbb{R}$.

If in addition, $s \neq \mathbb{R}^n$, we say that S is a proper subspace.

Definition 7 (Subspace)

A nonempty subset S of $\mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if $ax + by \in S$ for every $x, y \in S$ and every $a, b \in \mathbb{R}$.

If in addition, $s \neq \mathbb{R}^n$, we say that S is a proper subspace.

Definition 8 (Span)

The span of a finite number of vectors $x^1, \ldots, x^K$ is the subspace of $\mathbb{R}^n$ defined as the set of vectors $y = \sum_{k=1}^{K} a_k x^k$, where each $a_k \in \mathbb{R}$.

Any such y is called a linear combination of $x^1, \ldots, x^K$.

Definition 7 (Subspace)

A nonempty subset S of $\mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if $ax + by \in S$ for every $x, y \in S$ and every $a, b \in \mathbb{R}$.

If in addition, $s \neq \mathbb{R}^n$, we say that S is a proper subspace.

Definition 8 (Span)

The span of a finite number of vectors $x^1, \ldots, x^K$ is the subspace of $\mathbb{R}^n$ defined as the set of vectors $y = \sum_{k=1}^{K} a_k x^k$, where each $a_k \in \mathbb{R}$.

Any such y is called a linear combination of $x^1, \ldots, x^K$.

Definition 9 (Basis)

Given a subspace S of $\mathbb{R}^n$, with $S \neq \{0\}$, a basis of S is a collection of vectors that are linearly independent and whose span is equal to S.

Every basis of a given subspace has the same number of vectors and this number is called the dimension of the subspace.

Definition 7 (Subspace)

A nonempty subset S of $\mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if $ax + by \in S$ for every $x, y \in S$ and every $a, b \in \mathbb{R}$.

If in addition, $s \neq \mathbb{R}^n$, we say that S is a proper subspace.

Definition 8 (Span)

The span of a finite number of vectors $x^1, \ldots, x^K$ is the subspace of $\mathbb{R}^n$ defined as the set of vectors $y = \sum_{k=1}^{K} a_k x^k$, where each $a_k \in \mathbb{R}$.

Any such y is called a linear combination of $x^1, \ldots, x^K$.

Definition 9 (Basis)

Given a subspace S of $\mathbb{R}^n$, with $S \neq \{0\}$, a basis of S is a collection of vectors that are linearly independent and whose span is equal to S.

Every basis of a given subspace has the same number of vectors and this number is called the dimension of the subspace.

NOTE: If S is a proper subspace of $\mathbb{R}^n$, then there exists a nonzero vector a which is orthogonal to S, that is, $a^\intercal x = 0$ for every $x \in S$.

Definition 7 (Subspace)

A nonempty subset S of $\mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if $ax + by \in S$ for every $x, y \in S$ and every $a, b \in \mathbb{R}$.

If in addition, $s \neq \mathbb{R}^n$, we say that S is a proper subspace.

Definition 8 (Span)

The span of a finite number of vectors $x^1, \ldots, x^K$ is the subspace of $\mathbb{R}^n$ defined as the set of vectors $y = \sum_{k=1}^{K} a_k x^k$, where each $a_k \in \mathbb{R}$.

Any such y is called a linear combination of $x^1, \ldots, x^K$.

Definition 9 (Basis)

Given a subspace S of $\mathbb{R}^n$, with $S \neq \{0\}$, a basis of S is a collection of vectors that are linearly independent and whose span is equal to S.

Every basis of a given subspace has the same number of vectors and this number is called the dimension of the subspace.

NOTE: If S is a proper subspace of $\mathbb{R}^n$, then there exists a nonzero vector a which is orthogonal to S, that is, $a^T x = 0$ for every $x \in S$. If S has dimension $m < n$, there exist $n - m$ linearly independent vectors that are orthogonal to S.

Theorem 10

Suppose that the span S of the vectors $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^K$ has dimension m. Then:

1. *There exists a basis of S consisting of m of the vectors $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^K$.*
2. *If $k \leq m$ and $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k$ are linearly independent, we can form a basis of S by starting with $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k$, and choosing $m - k$ of the vectors $\boldsymbol{x}^{k+1}, \ldots, \boldsymbol{x}^K$.*

(Sketch of Proof):

Affine subspaces

Let S_0 be a subspace of $\mathbb{R}^n$ and let $\boldsymbol{x}^0$ be some vector. If we add $\boldsymbol{x}^0$ to every element of S_0, we are translating S_0 by $\boldsymbol{x}^0$, obtaining the affine subspace:

$$S = S_0 + \boldsymbol{x}^0 = \{\boldsymbol{x} + \boldsymbol{x}^0 \mid \boldsymbol{x} \in S_0\}.$$

In general, S is not a subspace, because it does not necessarily contain $\mathbf{0}$.
The dimension of S is defined to be equal to the dimension of S_0.

Affine subspaces

Let S_0 be a subspace of $\mathbb{R}^n$ and let $\boldsymbol{x}^0$ be some vector. If we add $\boldsymbol{x}^0$ to every element of S_0, we are translating S_0 by $\boldsymbol{x}^0$, obtaining the affine subspace:

$$S = S_0 + \boldsymbol{x}^0 = \{\boldsymbol{x} + \boldsymbol{x}^0 \mid \boldsymbol{x} \in S_0\}.$$

In general, S is not a subspace, because it does not necessarily contain $\mathbf{0}$.
The dimension of S is defined to be equal to the dimension of S_0.

Example: Affine subspace

Let $\boldsymbol{A}$ be an $m \times n$ matrix and $\boldsymbol{b} \in \mathbb{R}^m$, and consider the set

$$S = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\},$$

which we assume nonempty. Let us fix some $\boldsymbol{x}^0$ such that $\boldsymbol{A}\boldsymbol{x}^0 = \boldsymbol{b}$.
Hence, an arbitrary vector $\boldsymbol{x}$ belongs to S if and only if

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}^0 \qquad \Rightarrow \qquad \boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}^0) = \mathbf{0}.$$

Hence, we have $\boldsymbol{x} \in S$ if and only if $(\boldsymbol{x} - \boldsymbol{x}^0) \in S_0 = \{\boldsymbol{y} \mid \boldsymbol{A}\boldsymbol{y} = \mathbf{0}\}$.
We conclude that $S = \{\boldsymbol{y} + \boldsymbol{x}^0 \mid \boldsymbol{y} \in S_0\}$ and S is an affine subspace of $\mathbb{R}^n$.

References for these slides

- Chap. 1 in Bertsimas, Dimitris, and John N. Tsitsiklis. **Introduction to Linear Optimization**, 1997. Athena Scientific.