COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Math Methods for Political Science

Lecture 2: Systems of Linear Equations

Thibault Vatter<tv2233@columbia.edu>
Department of Statistics, Columbia University

September 11, 2017

The basics

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

Definition 1 (Linear equation and system of linear equations)

For $n \in \mathbb{N}_+$, a **linear equation** in the variables $x_1, \cdots, x_n$ is an equation that can be written as

$$a_1 x_1 + \cdots + a_n x_n = b,$$

where $a_1, \cdots, a_n, b \in \mathbb{R}$.

A **system of linear equations** (or **linear system**) is a collection of one or more linear equations in the same variables $x_1, \cdots, x_n$.

Example 1 (System of linear equations)

$$\begin{aligned} x_1 \quad -2x_2 &= -1 \\ -x_1 \quad +3x_2 &= 3 \end{aligned}$$

Definition 2 (Solutions to system of linear equations)

$(s_1, \cdots, s_n)$ is a **solution** of a linear system in $n \in \mathbb{N}_+$ variables if it makes every equation in the system true when $s_1, \cdots, s_n$ are substituted for $x_1, \cdots, x_n$ respectively.

Example 2 (Solutions to system of linear equations)

$$\begin{aligned} x_1 \quad -2x_2 &= -1 \\ -x_1 \quad +3x_2 &= 3 \end{aligned}$$

If $s_1 = 3$ and $s_2 = 2$, then

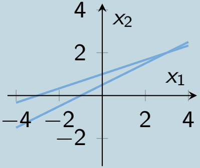
$$\begin{aligned} 3 \quad -2 \times 2 &= -1 \\ -3 \quad +3 \times 2 &= 3 \end{aligned}$$

Definition 3 (Solution set and equivalent systems)

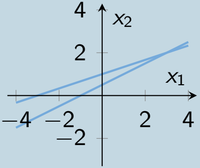
- The **solution set** of a linear system is the set of its solutions.
- Linear systems are **equivalent** if their solution sets are.

Example 3 (Solution set of 2d systems)

$$\begin{array}{rr} x_1 & -2x_2 = -1 \\ -x_1 & +3x_2 = 3 \end{array}$$



$$\begin{array}{rr} x_1/2 & -x_2 = -1/2 \\ -x_1/2 & +3/2x_2 = 3/2 \end{array}$$



Cardinality of the solution set

Lemma 1 (Cardinality of the solution set)

Let S be the solution set of a linear system, then $\mid S \mid \in \{0, 1, \infty\}$.

Definition 4 (Consistent linear system)

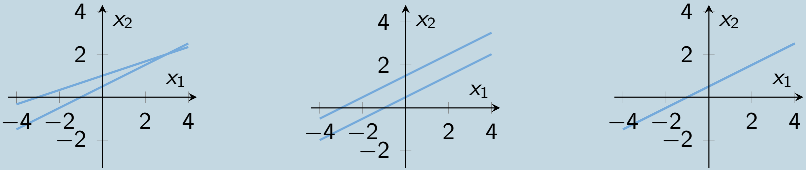
A system is **consistent** if its solution set S is s.t. $\mid S \mid \in \{1, \infty\}$.

Example 4 (Solution set of 2d systems cont'd)

$$\begin{aligned} x_1 \quad -2x_2 &= -1 \\ -x_1 \quad +3x_2 &= 3 \end{aligned}$$

$$\begin{aligned} x_1 \quad -2x_2 &= -1 \\ -x_1 \quad +2x_2 &= 3 \end{aligned}$$

$$\begin{aligned} x_1 \quad -2x_2 &= -1 \\ -x_1 \quad +2x_2 &= 1 \end{aligned}$$



Matrix

Definition 5 (Matrix)

For $n, d \in \mathbb{N}_+$, a $d \times n$ **matrix** is a rectangular array with d rows and n columns of numbers called **entries** (or **elements**).

Example 5 (Matrix)

$$\begin{bmatrix} 1 & 6 & 8 & 3 \\ 4 & 1000 & 3.14 & 9 \\ 2 & 5 & 0 & 0 \end{bmatrix}$$

For $n, d \in \mathbb{N}_+$ and A a $d \times n$ matrix, we write

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{1,n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{d,1} & A_{d,2} & \cdots & A_{d,n} \end{bmatrix}$$

Matrix notation for linear system

Definition 6 (Augmented matrix)

Consider the following linear system

$$\begin{array}{ccccccc} a_{1,1}x_1 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ \vdots & + & \vdots & + & \vdots & = & \vdots \\ a_{d,1}x_1 & + & \cdots & + & a_{d,n}x_n & = & b_d \end{array}$$

Then its **augmented matrix** is the $d \times (n+1)$ matrix A s.t.

$$A = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} & b_1 \\ \vdots & \vdots & \vdots & \vdots \\ a_{d,1} & \cdots & a_{d,n} & b_n \end{bmatrix}$$

Example 6 (Augmented matrix)

$$\begin{array}{rl} x_1 & -2x_2 = -1 \\ -x_1 & +3x_2 = 3 \end{array} \implies \text{the augmented matrix is } \begin{bmatrix} 1 & -2 & -1 \\ -1 & 3 & 3 \end{bmatrix}$$

Elementary row operations

Definition 7 (The three elementary row operations)

- **Replacement:** add to one row a multiple of another row.
- **Interchange:** interchange two rows.
- **Scaling:** multiply all entries in a row by a nonzero constant.

Example 7 (Row operations)

$$\begin{bmatrix} 1 & -2 & -1 \\ -1 & 3 & 3 \end{bmatrix} \Longrightarrow \begin{cases} \text{Replacement: } \begin{bmatrix} 1 & -2 & -1 \\ 0 & 1 & 2 \end{bmatrix} \\ \text{Interchange: } \begin{bmatrix} -1 & 3 & 3 \\ 1 & -2 & -1 \end{bmatrix} \\ \text{Scaling: } \begin{bmatrix} 1 & -2 & -1 \\ -2 & 6 & 6 \end{bmatrix} \end{cases}$$

Row equivalence

Definition 8 (Row equivalence)

Matrices A and B are **row equivalent** (r.e.) if there is a sequence of elementary row operations that transforms A into B.

Lemma 2 (Solution sets of r.e. augmented matrices)

Let S_1 and S_2 be the solution sets of linear systems with augmented matrices A_1 and A_2, then $S_1 = S_2 \Leftrightarrow A_1$ and A_2 are r.e.

Example 8 (Solution sets and augmented matrices)

$$\begin{bmatrix} 1 & -2 & -1 \\ -1 & 3 & 3 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & -2 & -1 \\ 0 & 1 & 2 \end{bmatrix} \Longrightarrow x_2 = 2, x_1 = 3$$

Echelon and reduced echelon forms

Definition 9 (Row echelon and reduced echelon forms)

A matrix is in **row echelon form** (**REF**) if

- nonzero rows are above any row of all zeros,
- the **leading entry** (the row's first non-zero element) is in a column to the right of the leading entry of the row above it.

A matrix is in **row reduced echelon form** (**RREF**) if it is in REF

- its nonzero leading entries are equal to 1,
- each leading 1 is the only nonzero entry in its column.

Example 9 (Row echelon and reduced echelon forms)

$$\text{REF: } \begin{bmatrix} 1 & -2 & -1 \\ 0 & 3 & 3 \\ 0 & 0 & 0 \end{bmatrix}, \text{ RREF: } \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

Pivots and row reduction algorithm

Definition 10 (Pivot position and pivot column)

A **pivot position** is a position of a leading 1 in a matrix in RREF, and a **pivot column** is a column containing a pivot position.

Definition 11 (The row reduction algorithm)

1. Start with the leftmost nonzero column.
2. Select a nonzero entry in this column as pivot. If necessary, interchange rows to move this entry at the top.
3. Use row operations to create zeros in all positions below.
4. Ignore the row with the pivot, apply 1-3 to the remaining submatrix, repeat until there are no more nonzero rows.
5. Beginning with the rightmost pivot and working upward, create zeros above each pivot. If a pivot is not 1, use scaling.

Pivots and row reduction algorithm

Example 10 (Row reduction algorithm, steps 1-4)

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix} \Longrightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

Example 11 (Row reduction algorithm, step 5)

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \Longrightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

Theorem 1 (Existence and uniqueness of the RREF)

For each matrix A, there exists a unique RREF U that is r.e. to A.

Let A be a linear system's augmented matrix and U its RREF.

Definition 12 (Basic and free variable)

- **Basic variables** correspond to U's pivot columns.
- **Free variables** are the other variables.

Theorem 2 (Existence and uniqueness of a system's solution)

It is consistent $\Leftrightarrow$ an REF of A has no row of the form

$$\begin{bmatrix} 0 & \cdots & 0 & b \end{bmatrix},\ b \neq 0.$$

If the system is consistent, then it has a unique solution when there are no free variables and infinitely many solutions otherwise.

Solutions to linear systems cont'd

Definition 13 (Using row reduction to solve a linear system)

1. Write the augmented matrix.
2. Compute the echelon form to decide if the system is consistent.
   - If there is no solution, stop.
   - Otherwise, go to next step.
3. Obtain the RREF.
4. Write the system corresponding to the RREF.
5. Rewrite each nonzero equation with its basic variable expressed in terms of its free variables.

Solutions to linear systems cont'd

Example 11 (Using row reduction to solve a linear system)

$$\begin{array}{rrrrrcr} & 3x_2 & -6x_3 & +6x_4 & +4x_5 & = & -5 \\ 3x_1 & -7x_2 & +8x_3 & -5x_4 & +8x_5 & = & 9 \\ 3x_1 & -9x_2 & +12x_3 & -9x_4 & +6x_5 & = & 15 \end{array}$$

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix} \Longrightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

$\Longrightarrow$ consistent and with free variables

$$\begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \Longrightarrow \begin{array}{rcrrr} x_1 & = & -24 & +2x_3 & -3x_4 \\ x_2 & = & -7 & +2x_3 & -2x_4 \\ x_5 & = & 4 & & \end{array}$$

Vectors in $\mathbb{R}^n$

Definition 14 ($\mathbb{R}^n$)

For $n \in \mathbb{N}_+$, $\mathbb{R}^n$ is the set of $n \times 1$ matrices with real entries and its elements are called **vectors**.

We denote

- vectors in bold (e.g. $\mathbf{u} \in \mathbb{R}^n$),
- their entries in light and dropping the column (e.g., $u_i \equiv \mathbf{u}_{i,1}$)

Definition 15 (Vector equality, sum and scalar multiple)

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and $r \in \mathbb{R}$, then

- $\mathbf{u} = \mathbf{v} \iff u_i = v_i \ \forall i$ (vector **equality**).
- $\mathbf{z} = \mathbf{u} + \mathbf{v} \iff \mathbf{z} \in \mathbb{R}^n \,\&\, z_i = u_i + v_i \ \forall i$ (vector **sum**).
- $\mathbf{z} = r\mathbf{u} \iff \mathbf{z} \in \mathbb{R}^n \,\&\, z_i = r u_i \ \forall i$ (**scalar multiple**).

Example 12 (Vector sum and scalar multiple)

Let $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, and $r = 5$, then

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix} \text{ and } r\mathbf{u} = 5 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$$

Lemma 3 (Properties of vector sums and scalar multiples)

With $\mathbf{0}$ the zero vector, $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and $c, d \in \mathbb{R}$, we have

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
3. $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$
4. $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$
5. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{v} + c\mathbf{u}$
6. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$
7. $c(d\mathbf{u}) = (cd)\mathbf{u}$
8. $1\mathbf{u} = \mathbf{u}$

Linear combinations and spans

Columbia University in the City of New York

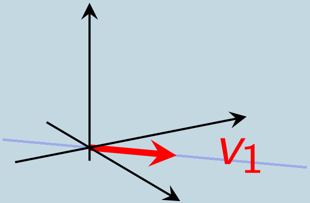
Definition 16 (Linear combinations)

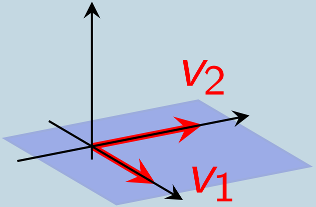
Given $\mathbf{v}_1, \cdots, \mathbf{v}_p \in \mathbb{R}^n$ and $c_1, \cdots, c_p \in \mathbb{R}$, $\mathbf{y} = \sum_{i=1}^{p} c_i \mathbf{v}_i$ is called a **linear combination** of $\mathbf{v}_1, \cdots, \mathbf{v}_p$ with **weights** $c_1, \cdots, c_p$.

Definition 17 (Spanned subsets of $\mathbb{R}^n$)

The subset of $\mathbb{R}^n$ spanned (or generated) by $\mathbf{v}_1, \cdots, \mathbf{v}_p \in \mathbb{R}^n$
Span $\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ is the set of linear combinations of $\mathbf{v}_1, \cdots, \mathbf{v}_p$.

Example 13 (Spanned subsets of $\mathbb{R}^3$)





The matrix vector product

Definition 18 (The matrix vector product)

Let A be a $m \times n$ matrix with columns $\mathbf{a}_1, \cdots, \mathbf{a}_n$ and $\mathbf{x} \in \mathbb{R}^n$. $A\mathbf{x} \in \mathbb{R}^m$ is the linear combination of the columns of A with the corresponding entries of $\mathbf{x}$ as weights, namely $A\mathbf{x} = \sum_{i=1}^n x_i \mathbf{a}_i$.

Example 14 (The matrix vector product for $m = n = 2$)

Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, then $A\mathbf{x} = 1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 2 \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$

Theorem 3 (Properties of the matrix vector product)

Let A be a $m \times n$ matrix, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and $c \in \mathbb{R}$, then

1. $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$,
2. $A(c\mathbf{u}) = c(A\mathbf{u})$.

The matrix equation $A\mathbf{x} = \mathbf{b}$

Let A be a $m \times n$ matrix with columns $\mathbf{a}_1, \cdots, \mathbf{a}_n$ and $\mathbf{x} \in \mathbb{R}^n$.

Theorem 4 (The matrix equation)

Let $\mathbf{b} \in \mathbb{R}^m$, then $A\mathbf{x} = \mathbf{b}$ has the same solution set as the linear system with augmented matrix $\begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n & \mathbf{b} \end{bmatrix}$.

Note that $A\mathbf{x} = \mathbf{b}$ has a solution $\iff$ $\mathbf{b} \in \text{Span}\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$.

Theorem 5 (Solutions of the matrix equation)

The following statements are equivalent:

1. *For each $\mathbf{b} \in \mathbb{R}^m$, then $A\mathbf{x} = \mathbf{b}$ has (at least) one solution.*
2. *Each $\mathbf{b} \in \mathbb{R}^m$ is a linear combination of the columns of A.*
3. *Span$\{\mathbf{a}_1, \cdots, \mathbf{a}_n\} = \mathbb{R}^m$.*
4. *A has a pivot position in every row.*