An elementary 3×2 system

- The following is an example of a system of 3 equations in 2 unknowns:

$$\begin{aligned} x_1 + 4.5x_2 &= 1, \\ 2x_1 + 1.2x_2 &= -3.2, \\ -0.1x_1 + 8.2x_2 &= 1.5. \end{aligned}$$

- This system can be written in vector format as $Ax = y$, where A is a 3×2 matrix, and y is a 3-vector:

$$A = \begin{bmatrix} 1 & 4.5 \\ 2 & 1.2 \\ -0.1 & 8.2 \end{bmatrix}, \quad y = \begin{bmatrix} 1 \\ -3.2 \\ 1.5 \end{bmatrix}.$$

- A solution to the linear equations is a vector $x \in \mathbb{R}^2$ that satisfies the equations.
- In the present example, it can be readily verified by hand calculations that the equations have no solution, i.e., the system is *infeasible*.

Example: polynomial interpolation

- Consider the problem of interpolating a given set of points (x_i, y_i), $i = 1, \ldots, m$, with a polynomial of degree $n-1$

$$p(x) = a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

- The polynomial interpolates the i-th point if and only if $p(x_i) = y_i$, and each of such conditions is a linear equation on the polynomial coefficients a_j, $j = 0, \ldots, n-1$.
- An interpolating polynomial is hence found if the following system of linear equations in the a_j variables has a solution:

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & & \cdots & \cdots & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^{n-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix},$$

where the matrix of coefficients on the left has a so-called Vandermonde structure.

Set of solutions of linear equations

- Generic linear equations can be expressed in vector format as

$$Ax = y,$$

where $x \in \mathbb{R}^n$ is the vector of unknowns, $y \in \mathbb{R}^m$ is a given vector, and $A \in \mathbb{R}^{m,n}$ is a matrix containing the coefficients of the linear equations.

- Key issues are: existence, uniqueness of solutions; characterization of the *solution set*:

$$S \doteq \{x \in \mathbb{R}^n : \; Ax = y\}.$$

- Let $a_1, \ldots, a_n \in \mathbb{R}^m$ denote the columns of A, i.e. $A = [a_1 \; \cdots \; a_n]$, and notice that the product Ax is nothing but a linear combination of the columns of A, with coefficients given by x:

$$Ax = x_1 a_1 + \cdots + x_n a_n.$$

- Ax always lies in $\mathcal{R}(A)$.
- Thus, $S \neq \emptyset \Leftrightarrow y \in \mathcal{R}(A)$.

Set of solutions of linear equations

- The linear equation
$$Ax = y, \quad A \in \mathbb{R}^{m,n}$$
admits a solution if and only if $\mathrm{rank}([A\ y]) = \mathrm{rank}(A)$.
- When this existence condition is satisfied, the set of all solutions is the affine set
$$S = \{x = \bar{x} + z : \ z \in \mathcal{N}(A)\},$$
where $\bar{x}$ is any vector such that $A\bar{x} = y$.
- In particular, the system has a unique solution if $\mathrm{rank}([A\ y]) = \mathrm{rank}(A)$ and $\mathcal{N}(A) = \{0\}$.

Overdetermined, underdetermined, and square systems

Overdetermined systems

- The system $Ax = y$ is said to be *overdetermined* when it has more equations than unknowns, i.e., when matrix A has more rows then columns ("skinny" matrix): $m > n$.
- Assume that A is full column rank, that is $\mathrm{rank}(A) = n$. Then, $\dim \mathcal{N}(A) = 0$, hence the system has either one or no solution at all.
- Indeed, the most common case for overdetermined systems is that $y \notin \mathcal{R}(A)$, so that no solution exists.
- In this case, it is often useful to introduce a notion of approximate solution, that is a solution that renders minimal some suitable measure of the mismatch between Ax and y (more on this later!)

Overdetermined, underdetermined, and square systems

Underdetermined systems

- The system $Ax = y$ is said to be *underdetermined* if it has more unknowns than equations, i.e., when matrix A has more columns than rows ("wide" matrix): $n > m$.
- Assume that A is full row rank, that is $\mathrm{rank}(A) = m$, and then $\mathcal{R}(A) = \mathbb{R}^m$, thus $\dim \mathcal{N}(A) = n - m > 0$.
- The system of linear equations is therefore solvable with infinite possible solutions, and the set of solutions has "dimension" $n - m$.
- Among all possible solutions, it is often of interest to single out one specific solution having minimum norm (more on this later!)

Overdetermined, underdetermined, and square systems

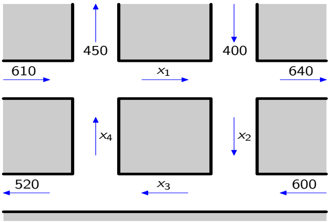
Square systems

- The system $Ax = y$ is said to be *square* when the number of equations is equal to the number of unknowns, i.e. when matrix A is square: $m = n$.
- If a square matrix is full rank, then it is invertible, and the inverse A^{-1} is unique and has the property that $A^{-1}A = I$.
- In the case of square full rank A the solution of the linear system is thus unique and it is formally written as

$$x = A^{-1}y.$$

Examples of linear equations

Network flows: a traffic example



$x_1 = ?, \quad x_2 = ?, \quad x_3 = ?, \quad x_4 = ?$

A basic traffic flow estimation problem involves inferring the amount of cars going through links based on information on the amount of cars passing through neighboring links.

Examples

Traffic example: flow equations

At each intersection, the incoming traffic has to match the outgoing traffic:

$$
\begin{array}{lrcl}
\text{Intersection A:} & x_4 + 610 & = & x_1 + 450, \\
\text{Intersection B:} & x_1 + 400 & = & x_2 + 640, \\
\text{Intersection C:} & x_2 + 600 & = & x_3, \\
\text{Intersection D:} & x_3 & = & x_4 + 520.
\end{array}
$$

We can write this in matrix format: $Ax = y$, with

$$
A = \begin{pmatrix} -1 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}, \quad y = \begin{pmatrix} -160 \\ 240 \\ -600 \\ 520 \end{pmatrix}.
$$

The matrix A is nothing else than the incidence matrix associated with the graph that has the intersections as nodes and links as edges.