Fundamentals of linear algebra

Germán Gómez-Herrero

Advanced Human Neurophysiology
VU University Amsterdam

November 8, 2012

The fundamental problem of linear algebra
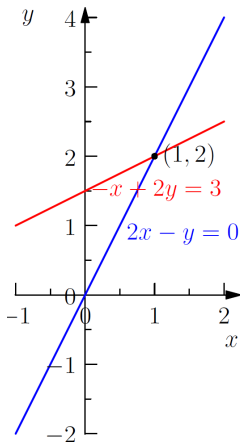
To solve n linear equations in n unknowns:

$$
\begin{aligned}
2x - y &= 0 \\
-x + 2y &= 3
\end{aligned}
$$

For $n = 2$, solving by **elimination** is easy:

$$
\left.
\begin{aligned}
2x - y &= 0 \quad \Rightarrow \quad y = 2x \\
-x + 2y &= 3
\end{aligned}
\right\} - x + 2(2x) = 3 \Rightarrow
\left\{
\begin{aligned}
x &= 1 \\
y &= 2
\end{aligned}
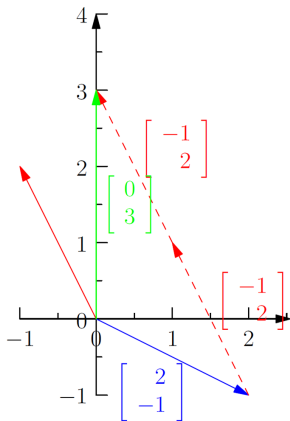\right.
$$

The geometry of linear equations
Row picture

- The set of all valid solutions to a linear equation with 2 unknowns defines a line in a 2-dimensional space

- The common solution to m linear equations must be in the intersection between the m lines that such equations define

The geometry of linear equations
Column picture

$$x \underbrace{\begin{bmatrix} 2 \\ -1 \end{bmatrix}}_{\mathbf{c}} + y \underbrace{\begin{bmatrix} -1 \\ 2 \end{bmatrix}}_{\mathbf{d}} = \underbrace{\begin{bmatrix} 0 \\ 3 \end{bmatrix}}_{v}$$

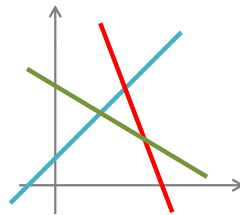In **vector** notation:

$$x\mathbf{c} + y\mathbf{d} = \mathbf{v}$$

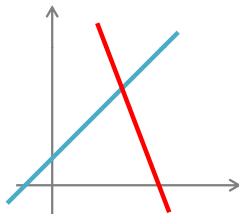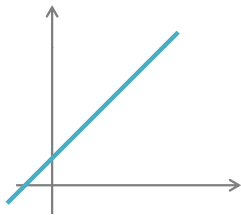Vector **v** is a **linear combination** of vectors **c** and **d**, and $x = -1$, $y = 2$ are the linear combination **coefficients**.

Linear equations
○○○●○

Linear independence
○○○○○○○○

Inverse
○○○○○

LS and MN solutions
○○○○○

Regularization
○○

References
○

A simple exercise

How many **unknowns** (aka variables, aka coefficients)?

How many **equations**?

How many **solutions**?

Linear equations
○○○●○
Linear independence
○○○○○○○○
Inverse
○○○○○
LS and MN solutions
○○○○○
Regularization
○○
References
○

A simple exercise

How many **unknowns** (aka variables, aka coefficients)?

How many **equations**?

How many **solutions**?

$$\#\text{vars} \;=\; 2$$
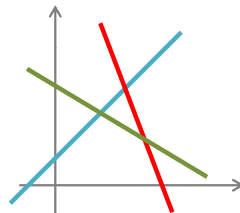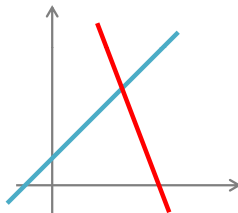$$\#\text{eqs} \;=\; 1$$
$$\#\text{sols} \;=\; \infty$$

A simple exercise

How many **unknowns** (aka variables, aka coefficients)?

How many **equations**?

How many **solutions**?

$$\#\text{vars} \quad = \quad 2 \qquad \#\text{vars} \quad = \quad 2$$
$$\#\text{eqs} \quad = \quad 1 \qquad \#\text{eqs} \quad = \quad 2$$
$$\#\text{sols} \quad = \quad \infty \qquad \#\text{sols} \quad = \quad 1$$
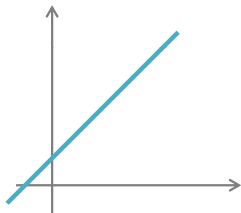
A simple exercise

How many **unknowns** (aka variables, aka coefficients)?

How many **equations**?

How many **solutions**?
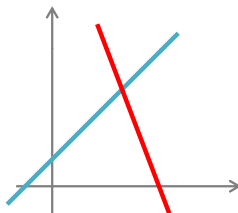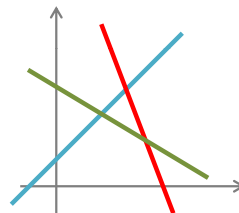
$\#\text{vars}$	$=$	2		$\#\text{vars}$	$=$	2		$\#\text{vars}$	$=$	2
$\#\text{eqs}$	$=$	1		$\#\text{eqs}$	$=$	2		$\#\text{eqs}$	$=$	3
$\#\text{sols}$	$=$	∞		$\#\text{sols}$	$=$	1		$\#\text{sols}$	$=$	0

A simple exercise
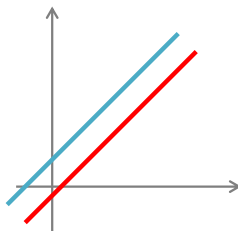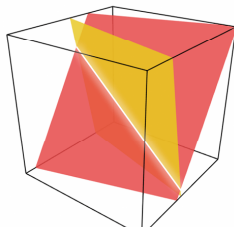
How many **unknowns** (aka variables)?

How many **equations**?

How many **solutions**?

A simple exercise

How many **unknowns** (aka variables)?

How many **equations**?
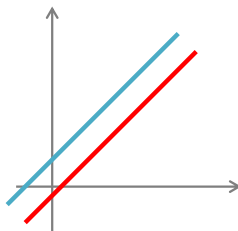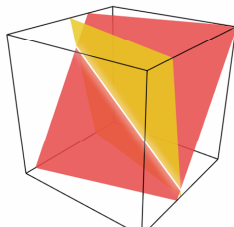
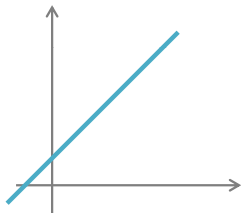How many **solutions**?

$$\#\text{vars} = 3$$
$$\#\text{eqs} = 2$$
$$\#\text{sols} = \infty$$

A simple exercise

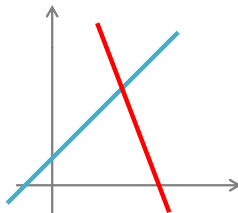How many **unknowns** (aka variables)?

How many **equations**?

How many **solutions**?
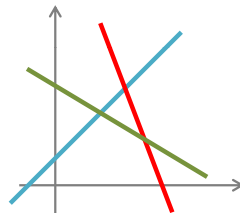

$$\begin{aligned} \#\text{vars} &= 3 \\ \#\text{eqs} &= 2 \\ \#\text{sols} &= \infty \end{aligned} \qquad \begin{aligned} \#\text{vars} &= 2 \\ \#\text{eqs} &= 2 \\ \#\text{sols} &= 0 \end{aligned}$$
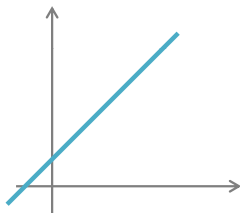
A simple exercise

How many **unknowns** (aka variables)?

How many **equations**?

How many **solutions**?
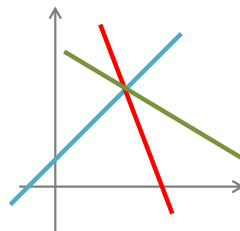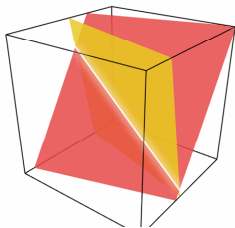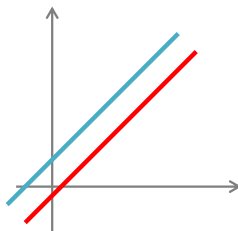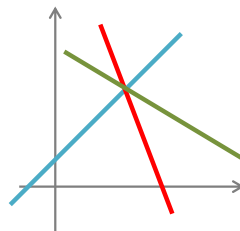

#vars	=	3	#vars	=	2	#vars	=	2
#eqs	=	2	#eqs	=	2	#eqs	=	3
#sols	=	∞	#sols	=	0	#sols	=	1

Elimination rapidly becomes cumbersome

What are the values of x_1, x_2, x_3, x_4, x_5, x_6, x_7?

$$
\begin{aligned}
-5x_1 + 2x_2 + 8x_3 + 4x_4 + 7x_5 - 6x_7 &= 4 \\
-x_1 + 3x_2 + 6x_3 - 7x_4 - 3x_5 + 8x_6 - 5x_7 &= 0 \\
9x_1 - 2x_2 - 8x_3 + 3x_4 + 4x_5 + 2x_6 + 7x_7 &= -1 \\
x_1 - 3x_2 - 5x_3 - 6x_5 + 2x_6 - 9x_7 &= -8 \\
9y - 3z + 5x_4 - 9x_5 + 7x_6 &= 3 \\
-5x_1 - 9x_2 + 3x_3 + 4x_4 + 5x_5 + 6x_6 - 6x_7 &= -9 \\
7x_2 - 7x_3 + 8x_4 + x_6 + 9x_7 &= -8
\end{aligned}
$$

Matrix picture

$$\underbrace{\begin{bmatrix} -5 & 2 & 8 & 4 & 7 & 0 & -6 \\ -1 & 3 & 6 & -7 & -3 & 8 & -5 \\ 9 & -2 & -8 & 3 & 4 & 2 & 7 \\ 1 & -3 & -5 & 0 & -6 & 2 & -9 \\ 0 & 9 & -3 & 5 & -9 & 7 & 0 \\ -5 & -9 & 3 & 4 & 5 & 6 & -6 \\ 0 & 7 & -7 & 8 & 0 & 1 & 9 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 4 \\ 0 \\ -1 \\ -8 \\ 3 \\ -9 \\ -8 \end{bmatrix}}_{\mathbf{b}}$$

In **matrix** notation:

$$\mathbf{Ax} = \mathbf{b}$$

Where **A** is a 7x7 **coefficient matrix**, **x** is 7x1, and **b** is 7x1.

Matrix multiplication

How do we multiply a matrix **A** by a vector **x**?

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = ?$$

Matrix multiplication

How do we multiply a matrix **A** by a vector **x**?

$$\left[\begin{array}{cc} 2 & 5 \\ 1 & 3 \end{array}\right] \left[\begin{array}{c} 1 \\ 2 \end{array}\right] =?$$

Method 1: take the **dot product** of each row of **A** with vector **x**:

$$\left[\begin{array}{cc} 2 & 5 \\ 1 & 3 \end{array}\right] \left[\begin{array}{c} 1 \\ 2 \end{array}\right] = \left[\begin{array}{ccc} 2 \cdot 1 & + & 5 \cdot 2 \\ 1 \cdot 1 & + & 3 \cdot 2 \end{array}\right] = \left[\begin{array}{c} 12 \\ 7 \end{array}\right]$$

Matrix multiplication

How do we multiply a matrix **A** by a vector **x**?

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = ?$$

Method 1: take the **dot product** of each row of **A** with vector **x**:

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \cdot 1 & + & 5 \cdot 2 \\ 1 \cdot 1 & + & 3 \cdot 2 \end{bmatrix} = \begin{bmatrix} 12 \\ 7 \end{bmatrix}$$

Method 2: think of the entries of **x** as the coefficients of a linear combination of two column vectors:

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 12 \\ 7 \end{bmatrix}$$

Linear independence

Recall our simple system of two linear equations:

$$\begin{array}{rcl} 2x - y & = & 0 \\ -x + 2y & = & 3 \end{array}$$

which can be written in matrix form:

$$\underbrace{\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} x \\ y \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ 3 \end{bmatrix}}_{\mathbf{b}}$$

Thinking x and y as coefficients in a linear combination of vectors:

$$x \underbrace{\begin{bmatrix} 2 \\ -1 \end{bmatrix}}_{\mathbf{a}_1} + y \underbrace{\begin{bmatrix} -1 \\ 2 \end{bmatrix}}_{\mathbf{a}_2} = \underbrace{\begin{bmatrix} 0 \\ 3 \end{bmatrix}}_{\mathbf{b}}$$

Linear independence

$$x \underbrace{\begin{bmatrix} 2 \\ -1 \end{bmatrix}}_{\mathbf{a}_1} + y \underbrace{\begin{bmatrix} -1 \\ 2 \end{bmatrix}}_{\mathbf{a}_2} = \underbrace{\begin{bmatrix} 0 \\ 3 \end{bmatrix}}_{\mathbf{b}}$$

We know that this system of equations has one unique solution.
But, is that true for any **b**?

Linear independence

$$x \underbrace{\left[\begin{array}{r} 2 \\ -1 \end{array}\right]}_{\mathbf{a}_1} + y \underbrace{\left[\begin{array}{r} -1 \\ 2 \end{array}\right]}_{\mathbf{a}_2} = \underbrace{\left[\begin{array}{c} 0 \\ 3 \end{array}\right]}_{\mathbf{b}}$$

We know that this system of equations has one unique solution. But, is that true for any $\mathbf{b}$?

YES because $\mathbf{a}_1$ and $\mathbf{a}_2$ are **linearly independent** and thus the linear combinations of $\mathbf{a}_1$ and $\mathbf{a}_2$ fill the whole 2-dimensional plane.

Linear independence

A set of vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ are **linearly independent** if any of them cannot be written as a linear combination of the other two, i.e.

$$\mathbf{v}_i \neq \alpha\mathbf{v}_j + \beta\mathbf{v}_k \qquad \forall\alpha \quad \forall\beta \quad \forall i \neq j \neq k$$

Linear independence

For any system of m linear equations with n unknowns:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

Then

A has $k = n < m$ LI columns $\Rightarrow$ Unique or no solution
A has $k = m = n$ LI columns $\Rightarrow$ Unique solution
A has $k = m < n$ LI columns $\Rightarrow$ ∞ solutions

A can have at most m linearly independent columns. Why?

Why should I care?

In Neuroimaging (and in many other fields of Science) you are rarely able to measure directly the processes of interest.

Instead, you often measure a linear combination of your hidden (or latent) variables of interest.

For instance, in EEG, we don't measure directly tiny neural currents in brain tissue, but the electric potentials that such currents generate at the scalp:

$$\mathbf{v} = \mathbf{A}\mathbf{x}$$

where $\mathbf{x}$ is an $n \times 1$ vector with current values at n brain locations, $\mathbf{v}$ is an $m \times 1$ vector of potentials at m scalp locations, and $\mathbf{A}$ is an $m \times n$ **leadfield matrix**.

Back to the large system of equations

$$
\underbrace{\begin{bmatrix}
-5 & 2 & 8 & 4 & 7 & 0 & -6 \\
-1 & 3 & 6 & -7 & -3 & 8 & -5 \\
9 & -2 & -8 & 3 & 4 & 2 & 7 \\
1 & -3 & -5 & 0 & -6 & 2 & -9 \\
0 & 9 & -3 & 5 & -9 & 7 & 0 \\
-5 & -9 & 3 & 4 & 5 & 6 & -6 \\
0 & 7 & -7 & 8 & 0 & 1 & 9
\end{bmatrix}}_{\textbf{A}}
\underbrace{\begin{bmatrix}
x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7
\end{bmatrix}}_{\textbf{x}}
=
\underbrace{\begin{bmatrix}
4 \\ 0 \\ -1 \\ -8 \\ 3 \\ -9 \\ -8
\end{bmatrix}}_{\textbf{b}}
$$

In **matrix** notation:

$$\textbf{Ax} = \textbf{b}$$

Back to the large system of equations

$$
\underbrace{\begin{bmatrix}
-5 & 2 & 8 & 4 & 7 & 0 & -6 \\
-1 & 3 & 6 & -7 & -3 & 8 & -5 \\
9 & -2 & -8 & 3 & 4 & 2 & 7 \\
1 & -3 & -5 & 0 & -6 & 2 & -9 \\
0 & 9 & -3 & 5 & -9 & 7 & 0 \\
-5 & -9 & 3 & 4 & 5 & 6 & -6 \\
0 & 7 & -7 & 8 & 0 & 1 & 9
\end{bmatrix}}_{\mathbf{A}}
\underbrace{\begin{bmatrix}
x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7
\end{bmatrix}}_{\mathbf{x}}
=
\underbrace{\begin{bmatrix}
4 \\ 0 \\ -1 \\ -8 \\ 3 \\ -9 \\ -8
\end{bmatrix}}_{\mathbf{b}}
$$

In **matrix** notation:

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

How do we get $\mathbf{x}$?

Inverse matrix

For any square matrix **A**, its inverse (if it exists) is defined as the matrix $\mathbf{A}^{-1}$ such that:

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

where **I** is the **identity matrix**, i.e. a matrix with ones in the diagonal and zeroes everywhere else. If **A** is 4×4 then:

$$\mathbf{I}_4 = \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right]$$

If $\mathbf{A}^{-1}$ exists, then it is unique.

If we knew $\mathbf{A}^{-1}$, how could we use it to solve $\mathbf{A}\mathbf{x} = \mathbf{b}$?

Invertible systems / Invertible matrices

Given a system of equations $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{A}$ square then:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

if $\mathbf{A}$ is invertible.

Notice that this holds for any $\mathbf{b}$!

What condition do you think $\mathbf{A}$ must fulfill to be invertible?

How do we know if a matrix is invertible?

Any non-singular matrix is invertible.

A square matrix is **singular** if not all of its columns are linearly independent. You can check this in MATLAB using:

`det(A)`

If the determinant of **A** is zero, then **A** is singular.

How do we calculate the inverse of a matrix?

There are several techniques. A common one is the so-called
Gauss-Jordan elimination.

In MATLAB you can simply do:

```
inv(A)
```

What about non-square matrices?

Recall:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$
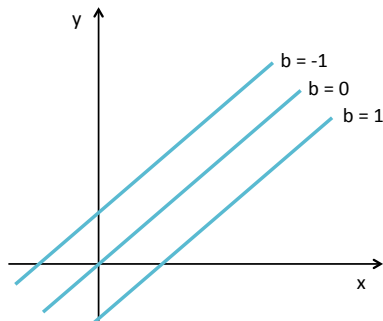
Then

A has $k = n < m$ LI columns $\Rightarrow$ Unique or no solution
A has $k = m < n$ LI columns $\Rightarrow$ ∞ solutions

Minimum norm solution for $m < n$

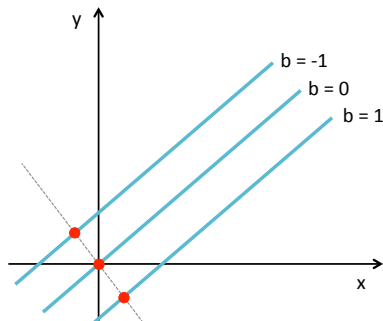An **underdetermined** system with $m = 1$, $n = 2$:

$$x - y = b \Rightarrow \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = b$$

Minimum norm solution for $m < n$

An **underdetermined** system with $m = 1$, $n = 2$:

$$x - y = b \Rightarrow \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = b$$

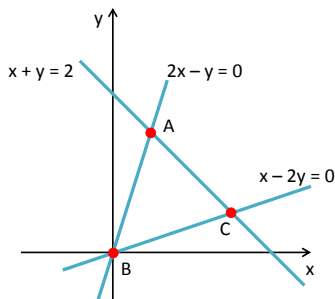Solution closest to the origin: The **minimum norm** solution.

Least squares solution for $m > n$

An **overdetermined** system with $m = 3$, $n = 2$:

$$
\begin{array}{rcl}
x + y & = & 2 \\
2x - y & = & 0 \\
x - 2y & = & 0
\end{array}
\Rightarrow
\left[\begin{array}{cc} 1 & 1 \\ 2 & -1 \\ 1 & -2 \end{array}\right]
\left[\begin{array}{c} x \\ y \end{array}\right]
=
\left[\begin{array}{c} 2 \\ 0 \\ 0 \end{array}\right]
$$

Least squares solution for $m > n$

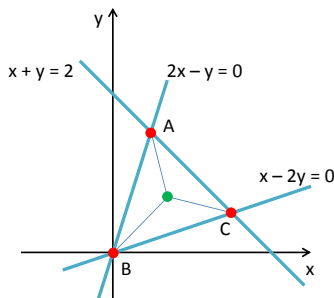An **overdetermined** system with $m = 3$, $n = 2$:

$$
\begin{array}{rcl}
x + y & = & 2 \\
2x - y & = & 0 \\
x - 2y & = & 0
\end{array}
\Rightarrow
\left[\begin{array}{rr} 1 & 1 \\ 2 & -1 \\ 1 & -2 \end{array} \right]
\left[\begin{array}{c} x \\ y \end{array} \right]
=
\left[\begin{array}{c} 2 \\ 0 \\ 0 \end{array} \right]
$$

Minimize Least Squared Error (LSE): The **least squares** solution.

Pseudoinverse

Given a linear system of m equations and n unknowns $\mathbf{Ax} = \mathbf{b}$ then, if the columns of $\mathbf{A}$ are linearly independent:

$$\hat{\mathbf{x}} = \mathbf{A}^+\mathbf{b}$$

with $\hat{\mathbf{x}}$ being:

- The Least Squares solution if $m > n$
- The Minimum Norm solution if $m < m$
- The exact solution if $m = n$

$\mathbf{A}^+$ is the **pseudoinverse** of $\mathbf{A}$, and can be easily obtained as:

$$\mathbf{A}^+ = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T$$

The pseudoinverse is just the inverse for $m = n$

For square matrices, the pseudoinverse reduces to the inverse:

$$\mathbf{A}^+ = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T = \mathbf{A}^{-1}\left(\mathbf{A}^T\right)^{-1}\mathbf{A}^T = \mathbf{A}^{-1}$$

In MATLAB: `pinv(A)`

But, what if $\mathbf{A}$ has linearly dependent columns?

Then neither $\mathbf{A}$ nor $\left(\mathbf{A}^T\mathbf{A}\right)$ are invertible, i.e. we cannot solve the linear system.

The pseudoinverse is just the inverse for $m = n$

For square matrices, the pseudoinverse reduces to the inverse:

$$\mathbf{A}^+ = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T = \mathbf{A}^{-1}\left(\mathbf{A}^T\right)^{-1}\mathbf{A}^T = \mathbf{A}^{-1}$$

In MATLAB: `pinv(A)`

But, what if $\mathbf{A}$ has linearly dependent columns?

Then neither $\mathbf{A}$ nor $\left(\mathbf{A}^T\mathbf{A}\right)$ are invertible, i.e. we cannot solve the linear system.

So, ... should we just give up?

Regularization

When **A** has linearly dependent columns the linear system is said
to be **ill-conditioned** and cannot be solved. However, we may still
find the solution of a **regularized** version of the original system of
linear equations. That is, instead of using:

$$\hat{\mathbf{x}} = \mathbf{A}^+\mathbf{b} \qquad \text{with} \quad \mathbf{A}^+ = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T$$

we are forced to use something like:

$$\hat{\mathbf{x}}_R = \mathbf{A}_{\mathbf{\Gamma}}^+\mathbf{b} \qquad \text{with} \quad \mathbf{A}_{\mathbf{\Gamma}}^+ = \left(\mathbf{A}^T\mathbf{A} + \mathbf{\Gamma}^T\mathbf{\Gamma}\right)^{-1}\mathbf{A}^T$$

for some suitably chosen regularization matrix **Γ**.

How does regularization affect the solution?

Regularization allows a numerical solution for **ill-conditioned** linear systems, at the expense of giving preference to certain solutions.

If regularization is based on valid a-priori information on the true solution, it enables a numerical solution of an otherwise unsolvable **inverse problem**.

If a too harsh regularization is used or if it is based on wrong a-priori information, it can lead to a numerically stable but completely wrong solution.

References and Attribution

These slides are partially based on course materials available at MIT Open Courseware:

http://ocw.mit.edu/courses/mathematics/
18-06sc-linear-algebra-fall-2011/

Some images are from Wikipedia:

http://en.wikipedia.org/wiki/System_of_linear_equations