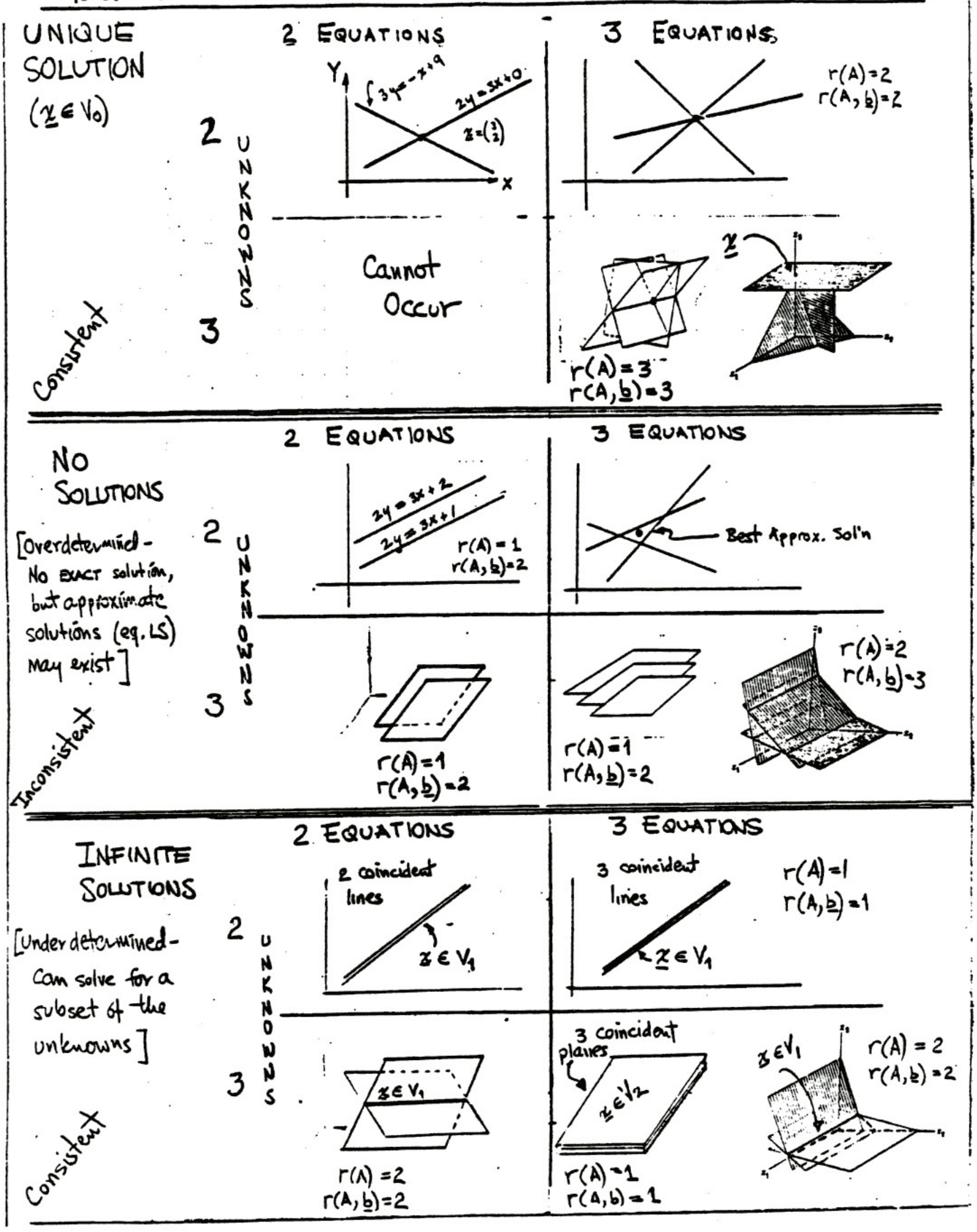
Solution of Non-Homogeneous SIMULTANEOUS EQUATIONS

Psychol. 614



UNIQUE SOLUTION ($\underline{x} \in V_0$)

Consistent

NO SOLUTIONS

[Overdetermined - No EXACT solution, but approximate solutions (eg. LS) may exist]

Inconsistent

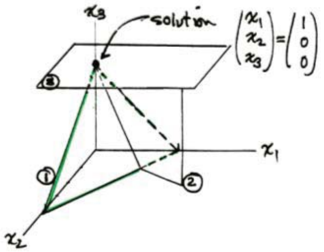
INFINITE SOLUTIONS

[Underdetermined - Can solve for a subset of the unknowns]

Consistent

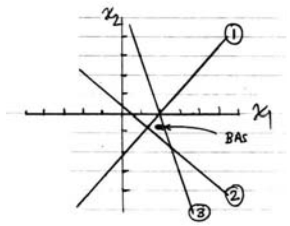
$$\underset{m\times n}{A}\ \underset{n\times 1}{\underline{x}} = \underset{m\times 1}{\underline{b}}$$

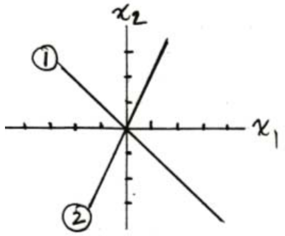
coefficients unknowns constants



Systems of linear equations

Michael Friendly
Psychology 6140





Linear equations in multivariate analysis

- Most problems in multivariate statistics involve solving a system of *m* equations in *n* unknowns

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n &= b_m \end{aligned}\right\} \longleftrightarrow \underset{m\times n}{A}\ \underset{n\times 1}{\underline{x}} = \underset{m\times 1}{\underline{b}}$$

coefficients unknowns constants

- When *m*=*n* and **A** is non-singular, solution is $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$
- It is useful to understand the general ideas, as well as the underlying geometry
- Counting unknowns (parameters) and independent equations (data) is important in understanding statistical models

---

Linear equations in multivariate analysis

- Two cases appear in statistical applications:
- **Non-homogeneous** equations:

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

  - The classic case is the general linear model, where we find estimates of regression coefficients and ANOVA effects by solving:

$$(\mathbf{X'X})\cdot\mathbf{b} = \mathbf{X'y}$$

- **Homogeneous** equations:

$$\mathbf{Ax} = \mathbf{0}$$

  - The classic case is in PCA/FA, where we find eigenvalues & eigenvectors by solving

$$(\mathbf{R} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$$

Linear equations: Examples

(A)

2 Unknowns

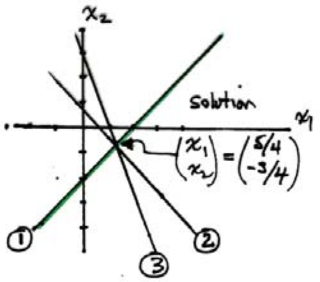
① $x_1 - x_2 = 2$
② $2x_1 + 2x_2 = 1$
③ $3x_1 + x_2 = 3$

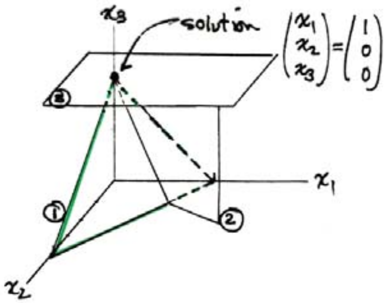
Each equation describes a <u>line</u> in 2D space

3 Unknowns

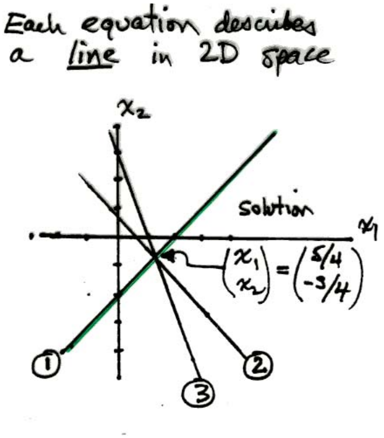
① $x_1 + x_2 + x_3 = 1$
② $x_1 - x_2 + 0 \cdot x_3 = 0$
③ $0 \cdot x_1 + 0 \cdot x_2 + x_3 = 1$

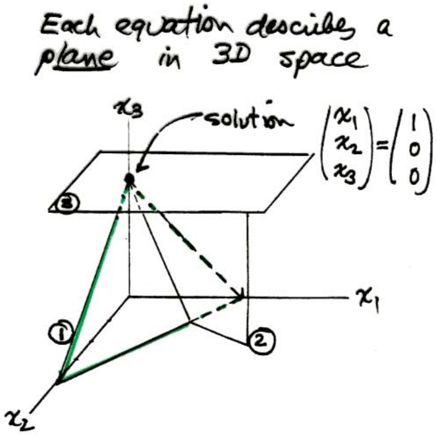
Each equation describes a <u>plane</u> in 3D space





Linear equations: Examples





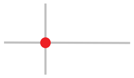
In these examples, we see that a solution (if one exists) corresponds to a **point** that lies on all three lines (in 2D) or in all three planes (3D) --- there they all **intersect** – thus satisfying **all** equations

---

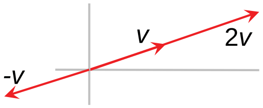
Vector space lingo

Def

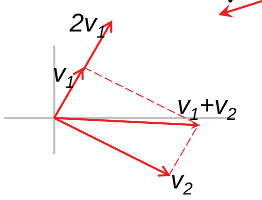
Def^n: A vector space, V_n, is the space of dimension *n* of all linear combinations of some set of vectors



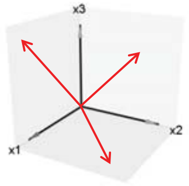
- V_0: v=(0, 0, 0, …) -- a point



- V_1: any one vector -- a line



- V_2: any two linearly independent vectors – a plane

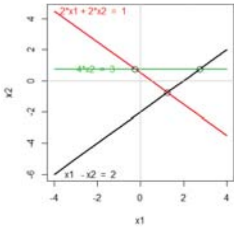


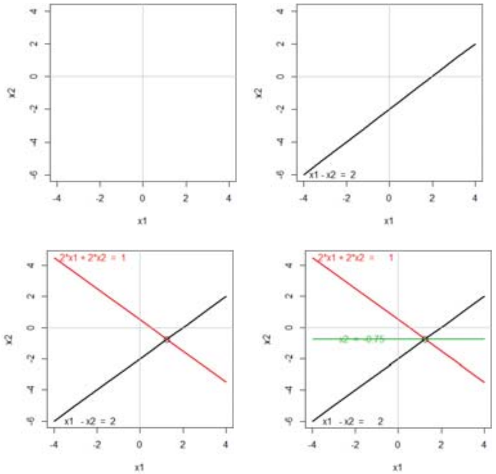
- V_3: any 3 linearly independent vectors – a volume

Geometric interpretation: 2D

Two unknowns

- Solution space is 2D (V_2)
- 1 eqn: 1D subspace (V_1)
- 2 eqn: 2 1D subspaces
  - May intersect in a point (V_0)
  - Unique solution
- 3 eqn: 3 1D subspaces
  - May intersect in a point (V_0)
  - Or, be inconsistent





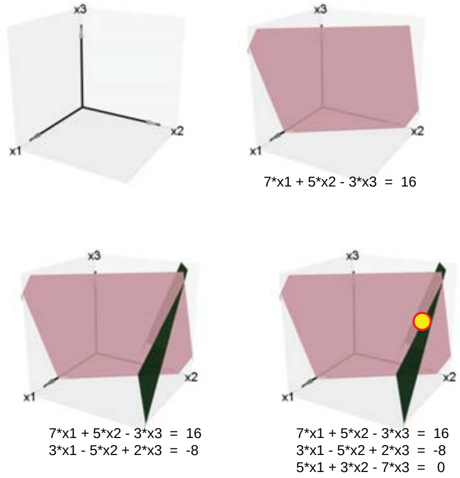
Geometric interpretation: 3D

Three unknowns

Solution space is 3D (V_3)

- 1 eqn: a 2D subspace (V_2)
  - a plane
- 2 eqn: 2 2D subspaces
  - may intersect in a line (V_1)
- 3 eqn: 3 2D subspaces
  - may intersect in a point (V_0) – unique solution

Each equation reduces solution space by 1 dimension (= 1 df)



Geometric interpretation of consistent equations

Generalizing: *n* unknowns → a space, V_n, of *n* dimensions

1 eqn in n unknowns → $(n-1)$D subspace, V_{n-1}, of V_n

2 eqn in n unknowns → 2 - V_{n-1} of V_n intersect in V_{n-2}

⋮

$(n-1)$ eqn in n unknowns → $(n-1)$ V_{n-1} of V_n intersect in $V_{n-(n-1)} = V_1$ = line

n eqn in n unknowns → n V_{n-1} of V_n intersect in $V_{n-n} = V_0$ = point

one unique solution

Not all systems have unique solutions

③

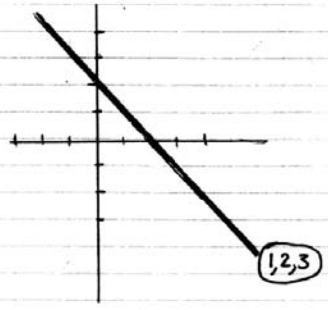
Underdetermined equations — Infinitely many solutions

Ⓑ

① $x_1 + x_2 = 2$

② $2x_1 + 2x_2 = 4$

③ $3x + 3x_2 = 6$



Here there are infinitely many pairs, x_1, x_2 which satisfy all 3 equations simultaneously — any pair of the form $(x_1, 2-x_1)$ —the equations are underdetermined (but consistent)

Note that we **can** find solutions by assigning an arbitrary value to one unknown, or adding one more equation ($x_2 = 0$). Then, $(x_1, x_2) = (2, 0)$ satisfies all equations.

Some systems have no (exact) solutions

④

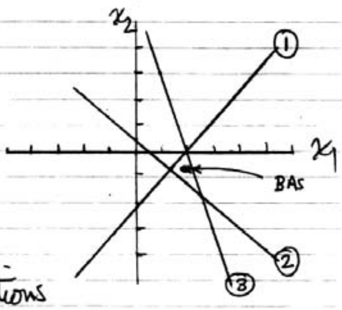
Overdetermined (inconsistent) equations — No Solution

Ⓒ

① $x_1 - x_2 = 2$

② $2x_1 + 2x_2 = 1$

③ $3x_1 + x_2 = 6$



Here, there is no pair of (x_1, x_2) values that simultaneously satisfy all 3 equations —the equations are inconsistent – or x_1, x_2 are overdetermined —but there is a solution for any 2 equations, or a best approximate solution (BAS) which comes 'closest' to all 3.

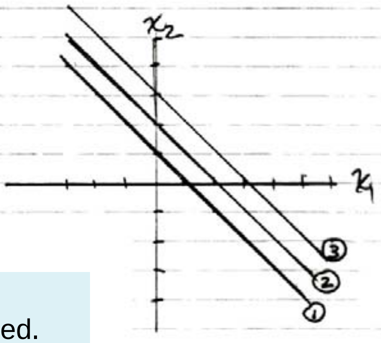
Some systems have even less!

Ⓓ

① $x_1 + x_2 = 1$

② $x_1 + x_2 = 2$

③ $x_1 + x_2 = 3$



These equations also have no solution (inconsistent) – x_1, x_2 are overdetermined.

Even any 2 equations are inconsistent.

Rank & geometry: conditions for solutions

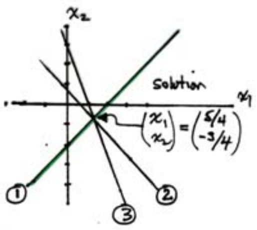
m equations in *n* unknowns: $\mathbf{A}_{m\times n}\mathbf{x}_{n\times 1} = \mathbf{b}_{m\times 1}$

- are ***consistent*** iff r(**A** | **b**) = r(**A**)
  - solution is ***unique*** if r(**A** | **b**) = r(**A**) = n
  - solution is ***underdetermined*** if r(**A** | **b**) = r(**A**) < n
- are ***inconsistent*** if r(**A** | **b**) < r(**A**)

Ⓐ Consistent/Unique

$$\underset{A}{\begin{bmatrix} 1 & -1 \\ 2 & 2 \\ 3 & 1 \end{bmatrix}} \underset{x}{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} = \underset{b}{\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}}$$

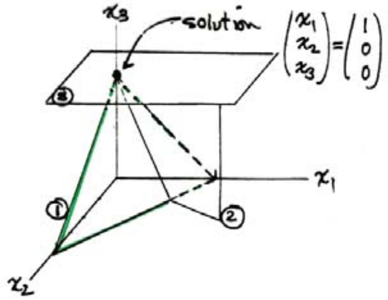
$r(A) = r(A|b) = 2$



Rank & geometry: conditions for solutions

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

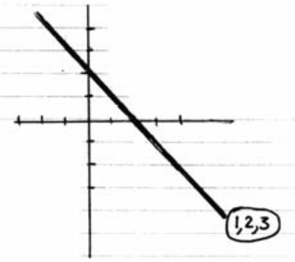
$r(A) = r(A|b) = 3$



Ⓑ Consistent/Underdetermined

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$$

$r(A) = r(A|b) = 1$



Rank & geometry: conditions for solutions

Inconsistent/Overdetermined

Ⓒ

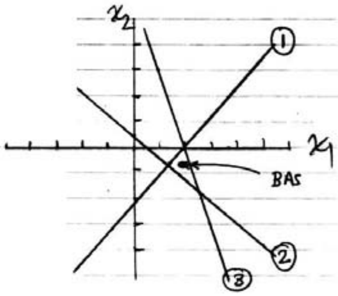
$$\begin{bmatrix} 1 & -1 \\ 2 & 2 \\ 3 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 6 \end{pmatrix}$$

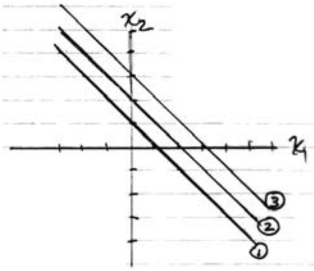
$r(A) = 2 \quad r(A|b) = 3$

Ⓓ

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

$r(A) = 1 \quad r(A|b) = 2$





Why r(A | b) = r(A) works

- Consistent / unique: r(**A** | **b**) = r(**A**) = n
  - All subspaces linearly independent, so they intersect in a unique point (V_0)
- / underdetermined: r(**A** | **b**) = r(**A**) < n
  - Some rows of **A** are linearly dependent, but the *same* dependence exists among elements of **b**
- Inconsistent: r(**A** | **b**) > r(**A**)
  - Linear relations among rows of **A** *differ* from those of **b** → there can be no (exact) solutions.

Statistical applications

(a) Regression

regression coefficients

$$\overset{n}{X}\,(\underline{b}) = (y)$$

predictors

subjects = m

– inconsistent — no exact solution unless m = n = # of regression coeffs

<u>But</u>, a least squares BAS is possible

Least squares principle → 'normal equations'

$$\underset{n\times n}{(X'X)}\ \underline{b} = \underset{n\times 1}{(X'y)}$$

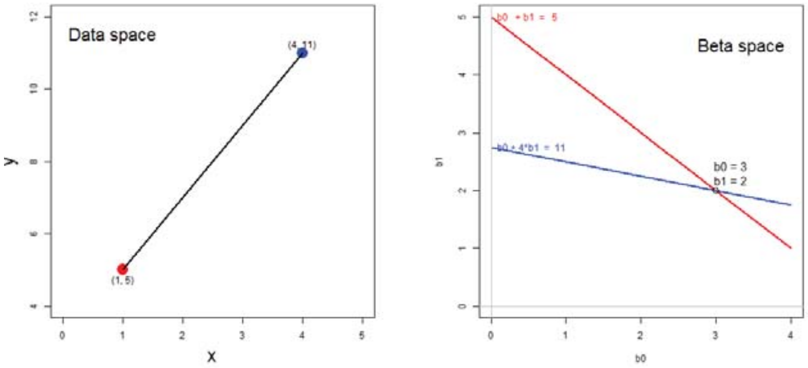
Since $r(X'X) = n$ (unless one predictor is a linear comb of others), solution is unique:

$$\underline{b} = (X'X)^{-1}X'y$$

Example: Find equation of a line through two (or more) points

- y = b0 + b1 x
- x = (1, 4), y = (5, 11)
- Each point provides one equation
- Solution for (b0, b1) satisfies all

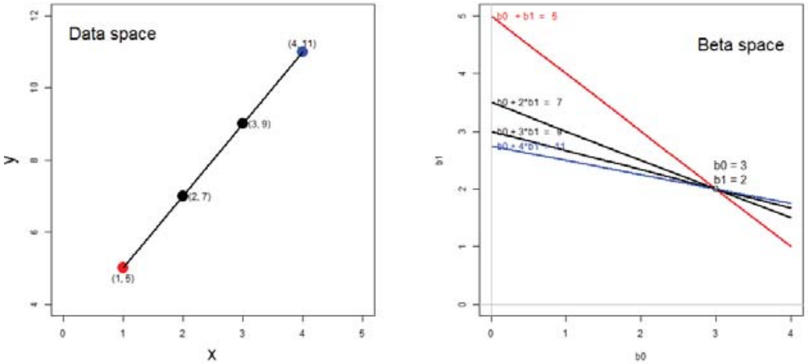
```
> showEqn(X, y)
b0   + b1  =   5
b0 + 4*b1  =  11
> Solve(X, y)
b0    =  3
  b1  =  2
```



Example: Find equation of a line through two (or more) points

- y = b0 + b1 x
- x = (1, 2, 3, 4), y = (5, 7, 9, 11)
- Each point provides one equation
- Solution for (b0, b1) satisfies all

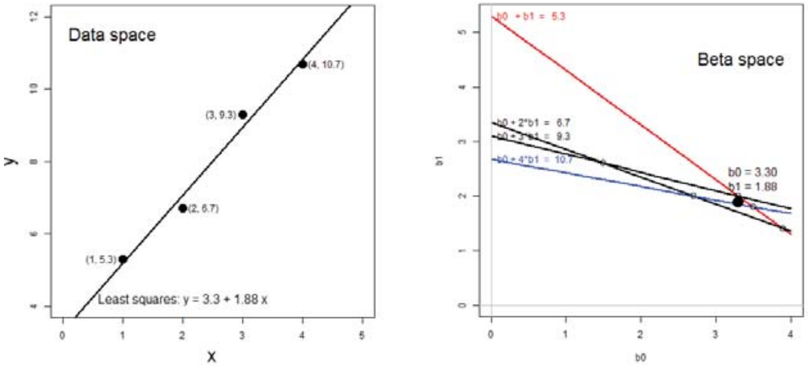
```
> showEqn(X, y)
b0   + b1  =   5
b0 + 2*b1  =   7
b0 + 3*b1  =   9
b0 + 4*b1  =  11
> Solve(X, y)
b0    =  3
  b1  =  2
   0  =  0
   0  =  0
```



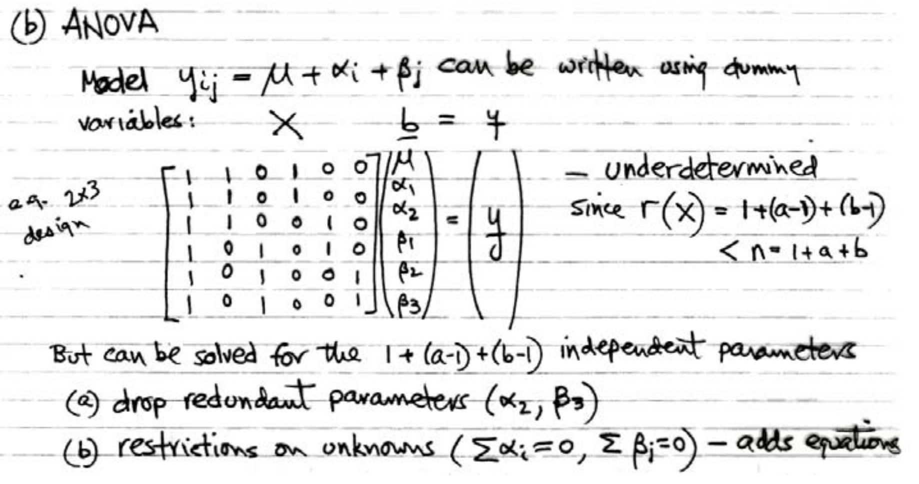
Real data never yields consistent equations

- y = b0 + b1 x
- x = (1, 2, 3, 4), y = (5.3 6.7 9.3 10.7)
- Each point provides one equation
- Least squares solution for (b0, b1) satisfies "best", in data space

```
> showEqn(X, y1)
b0   + b1  =   5.3
b0 + 2*b1  =   6.7
b0 + 3*b1  =   9.3
b0 + 4*b1  =  10.7
> Solve(X, y1)
x1    =   3.5
  x2  =   1.8
   0  =   0.4
   0  =  -0.4
```



Statistical applications



(b) ANOVA

Model $y_{ij} = \mu + \alpha_i + \beta_j$ can be written using dummy variables: X $\underline{b} = y$

e.g. 2x3 design

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} \\ \underline{y} \\ \\ \end{pmatrix}$$

– underdetermined since $r(X) = 1 + (a-1) + (b-1) < n = 1 + a + b$

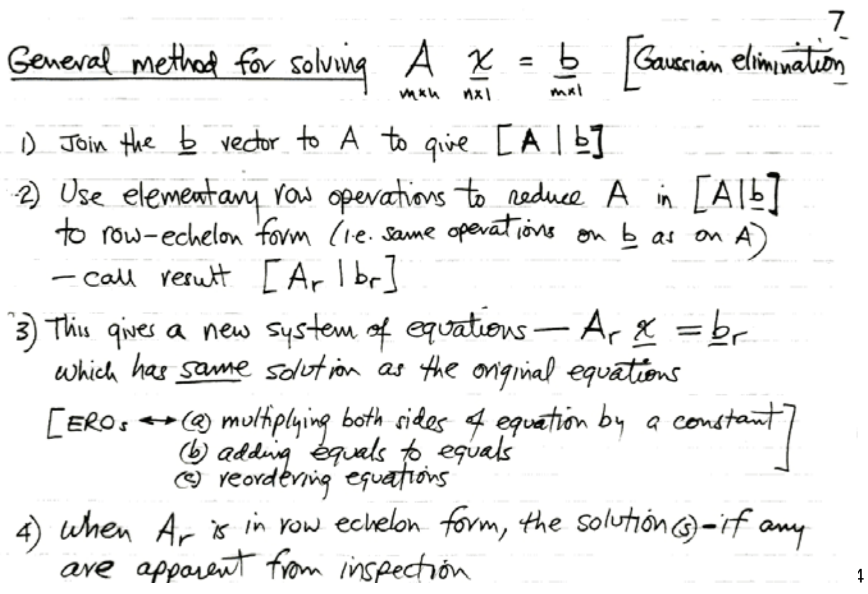
But can be solved for the $1 + (a-1) + (b-1)$ independent parameters

(a) drop redundant parameters (α_2, β_3)

(b) restrictions on unknowns $(\sum \alpha_i = 0, \sum \beta_j = 0)$ – adds equations

In general, factors and interactions in ANOVA designs have more parameters than can be solved for independently. Standard software (R, SAS, SPSS) handles this automatically, but **interpretation** depends on knowing what it does!

Gaussian elimination



7

General method for solving $\underset{m\times n}{A} \; \underset{n\times 1}{\underline{x}} = \underset{m\times 1}{\underline{b}}$ [Gaussian elimination]

1) Join the $\underline{b}$ vector to A to give $[A \mid \underline{b}]$

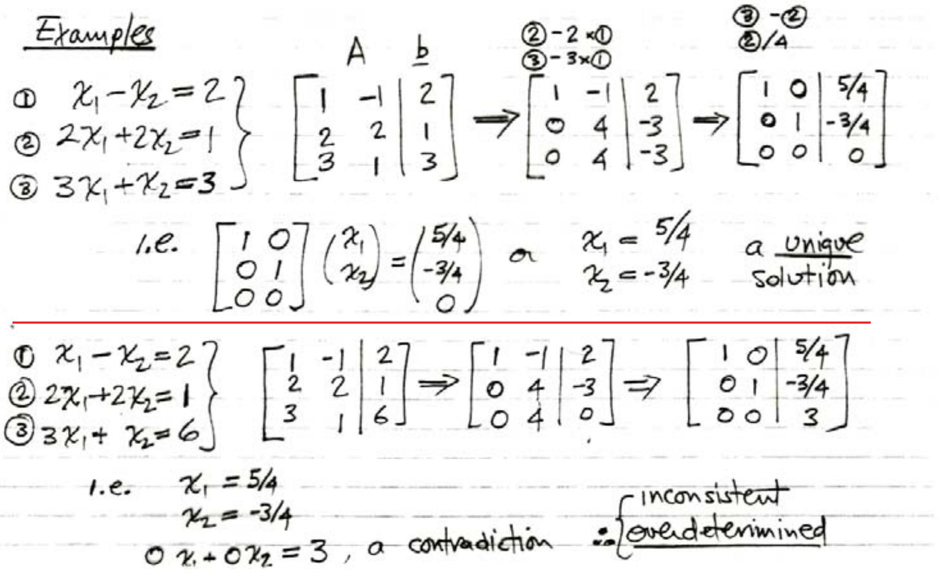
2) Use elementary row operations to reduce A in $[A \mid \underline{b}]$ to row-echelon form (i.e. same operations on $\underline{b}$ as on A) – call result $[A_r \mid b_r]$

3) This gives a new system of equations — $A_r \underline{x} = \underline{b}_r$ which has same solution as the original equations

[EROs ↔ (a) multiplying both sides of equation by a constant
(b) adding equals to equals
(c) reordering equations]

4) when A_r is in row echelon form, the solution(s) – if any are apparent from inspection

Gaussian elimination



Examples

① $x_1 - x_2 = 2$
② $2x_1 + 2x_2 = 1$
③ $3x_1 + x_2 = 3$

$$\begin{array}{c} A \quad\quad b \\ \left[\begin{array}{cc|c} 1 & -1 & 2 \\ 2 & 2 & 1 \\ 3 & 1 & 3 \end{array}\right] \end{array} \xrightarrow[③-3\times①]{②-2\times①} \left[\begin{array}{cc|c} 1 & -1 & 2 \\ 0 & 4 & -3 \\ 0 & 4 & -3 \end{array}\right] \xrightarrow[②/4]{③-②} \left[\begin{array}{cc|c} 1 & 0 & 5/4 \\ 0 & 1 & -3/4 \\ 0 & 0 & 0 \end{array}\right]$$

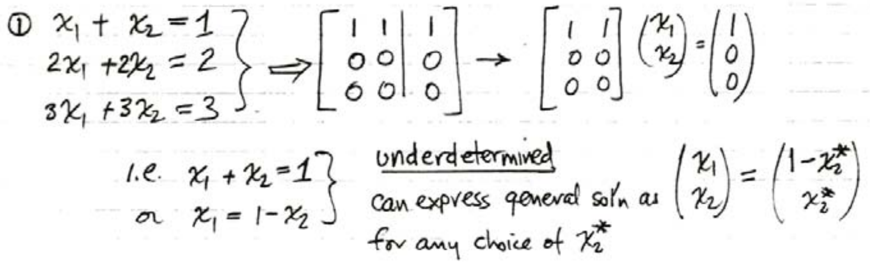
i.e. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5/4 \\ -3/4 \\ 0 \end{pmatrix}$ or $x_1 = 5/4$, $x_2 = -3/4$ a unique solution

① $x_1 - x_2 = 2$
② $2x_1 + 2x_2 = 1$
③ $3x_1 + x_2 = 6$

$$\left[\begin{array}{cc|c} 1 & -1 & 2 \\ 2 & 2 & 1 \\ 3 & 1 & 6 \end{array}\right] \Rightarrow \left[\begin{array}{cc|c} 1 & -1 & 2 \\ 0 & 4 & -3 \\ 0 & 4 & 0 \end{array}\right] \Rightarrow \left[\begin{array}{cc|c} 1 & 0 & 5/4 \\ 0 & 1 & -3/4 \\ 0 & 0 & 3 \end{array}\right]$$

i.e. $x_1 = 5/4$
$x_2 = -3/4$
$0x_1 + 0x_2 = 3$, a contradiction ∴ inconsistent / overdetermined

Gaussian elimination



① $x_1 + x_2 = 1$
$2x_1 + 2x_2 = 2$
$3x_1 + 3x_2 = 3$

$$\Rightarrow \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] \rightarrow \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

i.e. $x_1 + x_2 = 1$
or $x_1 = 1 - x_2$

underdetermined

can express general sol'n as $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 - x_2^* \\ x_2^* \end{pmatrix}$ for any choice of x_2^*

Gaussian elimination

General rules

- If any zero row of A_r corresponds to a non-zero entry of $\underline{b}_r$, then there is a contradiction & system is overdetermined
- If there is no contradiction & rank $A = n$ (= # non-zero rows in A_r) then solution is unique

e.g. for $n=3$, $m \geq 3$

$$\begin{array}{ccc} A_r & & b_r \\ \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\\0&0&0\\ &\vdots& \\0&0&0\end{bmatrix} & \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = & \begin{pmatrix}b_1\\b_2\\b_3\\0\\\vdots\\0\end{pmatrix}\end{array} \Rightarrow \begin{cases}x_1=b_1\\x_2=b_2\\x_3=b_3\end{cases}$$

If $m=n$ (equal # of equations & unknowns & if $r(A)=n$ then the unique solution can be expressed as $\underline{x} = A^{-1}\underline{b}$

Gaussian elimination

- If no contraction & rank $A = r < n$, then equations are underdetermined — $n-r$ unknowns can be given arbitrary values & remaining r unknowns can be solved in terms of these.

E.g.

$$\begin{array}{cccc} & A_r & & b_r & \\ & \begin{bmatrix}1&1&0&2\\0&0&1&-3\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} = & \begin{pmatrix}4\\-1\\0\\0\end{pmatrix} & \leftarrow \text{consistent}\end{array}$$

rank(A_r) = rank(A) = 2 → underdetermined, can solve for 2 unknowns

positions of leading 1s indicate which x's to solve for

i.e.

$$\left.\begin{array}{l} \textcircled{x_1} + x_2 + 2x_4 = 4 \\ \textcircled{x_3} - 3x_4 = -1 \end{array}\right\} \quad \begin{array}{l} x_1 = 4 - x_2^* - 2x_4^* \\ x_2 = x_2^* \\ x_3 = -1 + 3x_4^* \\ x_4 = x_4^* \end{array}$$

(x_2^*, x_4^* are arbitrary values)

Gaussian elimination- matrix algebra

An algebraic argument for the case $r(A|\underline{b}) = r(A) < n$ is as follows:

$r(A) = r < n \leftrightarrow$ some $m-r$ rows of A are linearly dependent on the remaining r rows

Permute the linearly independent r rows of A & b to come first and partition the equations:

$$\begin{array}{c} r\{ \\ m-r\{ \end{array}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{pmatrix}\underline{x}_1\\\underline{x}_2\end{pmatrix}\begin{array}{l}\}r\\\}n-r\end{array} = \begin{pmatrix}b_1\\b_2\end{pmatrix}\begin{array}{l}r\\m-r\end{array}$$

$$\underbrace{\qquad}_{r}\ \underbrace{\qquad}_{n-r}$$

Since the last $m-r$ rows (equations) are linearly dependent, they can be ignored, so we can solve the first r:

$$\begin{bmatrix} A_{11} & A_{12} \end{bmatrix}\begin{pmatrix}\underline{x}_1\\\underline{x}_2\end{pmatrix} = \underline{b}_1 \Rightarrow A_{11}\underline{x}_1 + A_{12}\underline{x}_2 = \underline{b}_1$$

$$\text{Solution} \begin{cases} \underline{x}_1 = A_{11}^{-1}\underline{b}_1 - A_{11}^{-1}A_{12}\underline{x}_2^* \\ x_2 = x_2^* \text{ (arbitrary)} \end{cases}$$

The matlib package

Functions for visualizing linear algebra and systems of equations:

- `showEqn(A, b)`
- `plotEqn(A, b)`
- `gaussianElimination(A, b)`
- `echelon(A, b)`
- `Solve(A, b)`

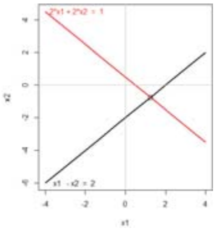
Many others, including:

- `R(A)` – matrix rank
- `tr(A)` – matrix trace

Install: `install.packages(“matlib”)`

Use: `library(matlib)`

```
> showEqn(A, b)
1*x1 - 1*x2  =  2
2*x1 + 2*x2  =  1
```

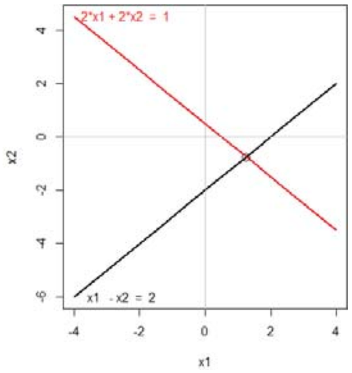


Two consistent equations in two unknowns:

```
> A <- matrix(c(1, 2, -1, 2), 2, 2)
> b <- c(2,1)
> showEqn(A, b)
1*x1 - 1*x2  =  2
2*x1 + 2*x2  =  1
> c( R(A), R(cbind(A,b)) )          # show ranks
[1] 2 2
> all.equal( R(A), R(cbind(A,b)) )  # consistent?
[1] TRUE
```

```
> plotEqn(A,b)

> Solve(A, b, fractions=TRUE)
x1    =   5/4
  x2  =  -3/4
```



```
> gaussianElimination(A, b, fractions=TRUE, verbose=TRUE)

Initial matrix:
     [,1] [,2] [,3]
[1,]  1   -1    2
[2,]  2    2    1

row: 1

 exchange rows 1 and 2
     [,1] [,2] [,3]
[1,]  2    2    1
[2,]  1   -1    2

 multiply row 1 by 1/2
     [,1] [,2] [,3]
[1,]   1    1  1/2
[2,]   1   -1    2
…

row: 2

 multiply row 2 by -1/2
     [,1] [,2] [,3]
[1,]    1    1  1/2
[2,]    0    1 -3/4

 subtract row 2 from row 1
     [,1] [,2] [,3]
[1,]    1    0  5/4
[2,]    0    1 -3/4
>
```

Homogeneous equations

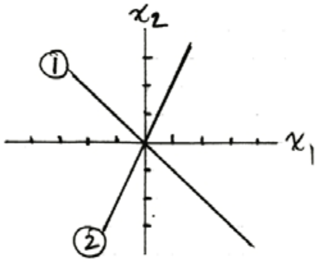
When the vector $\underline{b}$ is zero:

$$\underset{m\times n}{A}\ \underset{n\times 1}{\underline{x}} = \underset{m\times 1}{\underline{0}}$$

the system is called homogeneous

e.g. ① $x_1 - x_2 = 0$

② $x_1 + 2x_2 = 0$



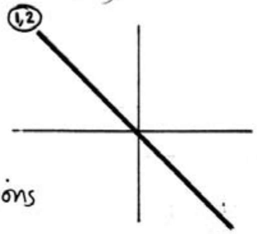
The trivial solution, $\underline{x} = 0$, always satisfies a homogeneous system — which therefore is never overdetermined

Homogeneous equations

Nontrivial solutions exist only when rank (A) < n — an underdetermined system

e.g. $x_1 - x_2 = 0$ ①

$2x_1 - 2x_2 = 0$ ②



Here, there are (infinitely many) solutions other than $\underline{x} = \underline{0}$

When r(A) < n, transformation to row echelon form can also be used to solve the system

e.g. $\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 2 & 2 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$ i.e. general solution is $\begin{cases} x_1 = -x_2^* \\ x_2 = \text{arbitrary} \end{cases}$

We will use this next week to help understand eigenvalues and eigenvectors

Summary

- Linear equations are used in solving for parameters in linear models
- Non-homogeneous equations: $\mathbf{A}\ \mathbf{x} = \mathbf{b}$
  - m = n: solution is $\mathbf{x} = \mathbf{A}^{-1}\ \mathbf{b}$ if $\mathbf{A}^{-1}$ exists
  - Consistent: if $r(\mathbf{A}\ |\ \mathbf{b}) = r(\mathbf{A})$
  - Inconsistent: if $r(\mathbf{A}\ |\ \mathbf{b}) > r(\mathbf{A})$
- Geometry:
  - Each *n*-variable eqn: (n-1) dim hyper plane in V_{n-1} subspace
  - Unique solution *iff* intersect in a point (V_0)
  - Underdetermined if intersect in larger space
  - Inconsistent if no points lie on all hyper planes