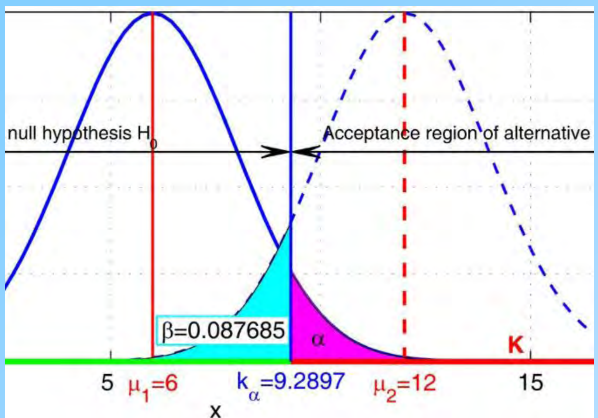
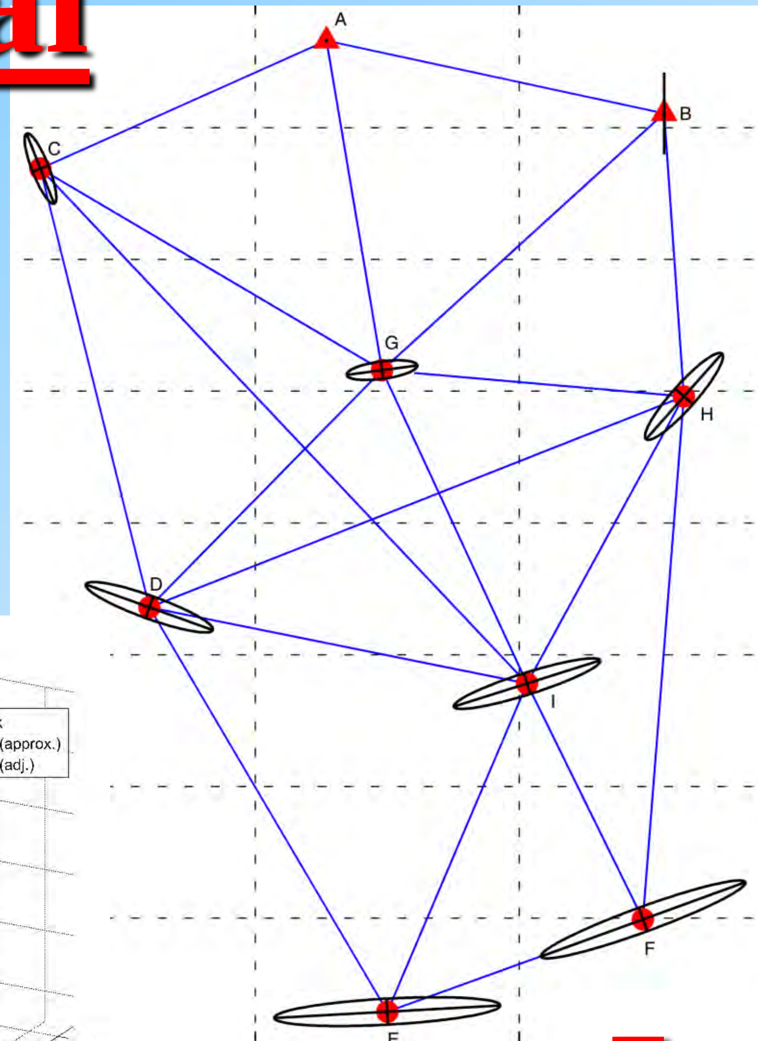
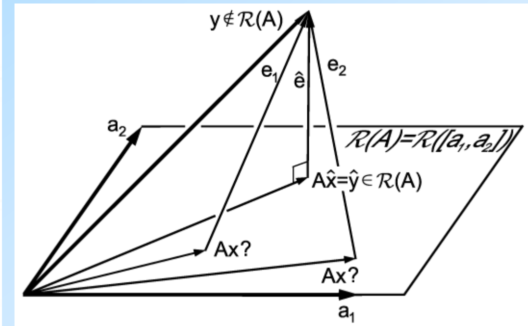
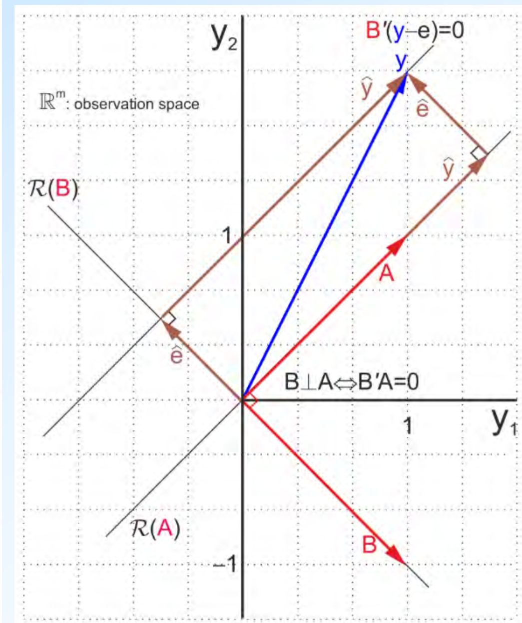
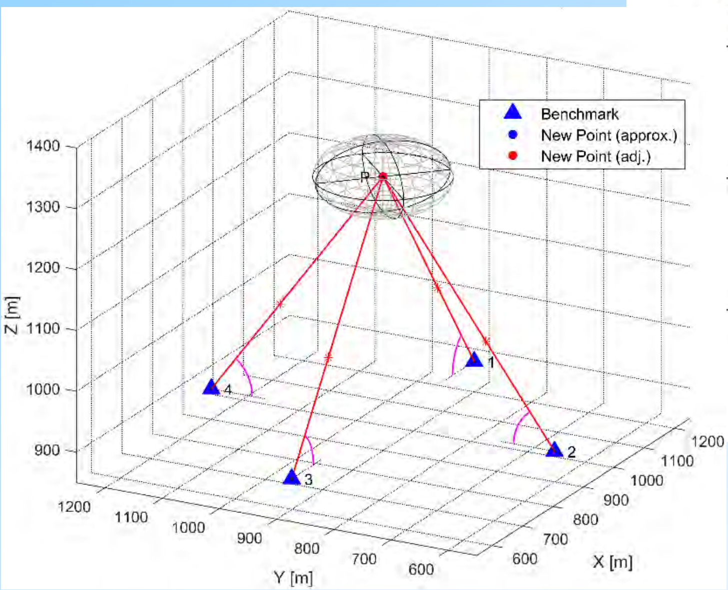
Statistical











Inference

Statistical Inference Rev. 2.44a

 GIS

Organization/Preliminary remarks

- Exam: 60 min, asap after winter term, together with "Signal Processing" (60 min) (no tools: "closed book"); assigned marks will be averaged with results of other module parts. Exam admission requirements: Acknowledged labs → Lab guidelines
- Lecture Notes incl. Numerical Examples (NE) : available from website
- Other material: Teunissen books ("Adjustment Theory", "Testing Theory" and "Network Quality Control"), can be borrowed from Institute's Library (5th floor) or University Library
- MATLAB stuff from website
- Got nothing to do ? Have too much spare time ? Want more work ? More explanations, more examples ? → Ask for a tutorial !

→ References

References

- Anderson J M and E M Mikhail (1998): Surveying. Theory and Practice. 7th edition. McGraw-Hill, ISBN 0-07-015914-9
- Ben-Israel A and T N E Greville (2001): Generalized Inverses: Theory and Applications.
- Benning W (2011): Statistik in Geodäsie, Geoinformation und Bauwesen. 4., überarbeitete und erweiterte Auflage, Wichmann, ISBN 978-3-87907-512-6
- Beucher O (2007): Wahrscheinlichkeitsrechnung und Statistik mit MATLAB. Springer, ISBN 978-3-540-72155-0
- Caspary W and K Wichmann (2007): Auswertung von Messdaten. Statistische Methoden für Geo- und Ingenieurwissenschaften. Oldenbourg, ISBN 978-3-486-58351-9
- Chatterjee S and A S Hadi (1988): Sensitivity Analysis in Linear Regression. John Wiley & Sons, Inc., ISBN 0-471-82216-7
- Chatterjee S and A S Hadi (2006): Regression Analysis by Example. 4th edition. John Wiley & Sons, Inc., ISBN 13 978-0-471-74696-6
- Ghilani Ch. D. (2010): Adjustment Computations. Spatial Data Analysis. 5th edition. John Wiley & Sons, Inc., ISBN 978-0-470-46491-5 (Errata Sheet)
- Grafarend E W (2006): Linear and Nonlinear Models - Fixed Effects, Random Effects, and Mixed Models. de Gruyter, ISBN 978-3-11-016216-5
- Jäger R et al. (2005): Klassische und robuste Ausgleichungsverfahren. Ein Leitfaden für Ausbildung und Praxis von Geodäten und Geoinformatikern. Wichmann, ISBN 3-87907-370-8
- Jäger R und R Bill (1986): Genauigkeits- und Zuverlässigkeitsmaße in Verdichtungsnetzen. Mitteilungen des DVW Baden-Württemberg 33, Sonderheft "Beurteilung geodätischer Netze".

→ References

References

- Koch K R (1999): Parameter Estimation and Hypothesis Testing in Linear Models. 2nd updated and enlarged edition. Springer, ISBN 978-3-540-65257-1
- Koch K R (2004): Parameterschätzung und Hypothesentests in linearen Modellen. Vierte, bearbeitete Auflage. Dümmlers
- Krumm F (2018): Geodetic Network Adjustment Examples
- Kuang S L (1996): Geodetic Network Analysis and Optimal Design: Concepts and Applications. Ann Arbor Press, Inc., Chelsea, Michigan, ISBN 1-57504-044-1
- Lay D C (2012): Linear Algebra and its Applications. 4th edition. Addison-Wesley Publishing Company, ISBN 978-0-321-38517-8
- Leick A et al. (2015): GPS Satellite Surveying, 4th edition. John Wiley & Sons, ISBN: 978-1-118-67557-1
- Magnus J R and H Neudecker (2007): Matrix Differential Calculus with Applications in Statistics and Econometrics. 3rd edition. John Wiley & Sons Ltd., ISBN 0-471-98632-1
- Mikhail E M and F Ackermann (1976): Observations and Least Squares. IEP-A Dun-Donnelley Publisher, ISBN 0-7002-2481-5
- MIT Course 18.06 on "Linear Algebra" by Prof. Gilbert Strang → http://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/
- Niemeier, W (2008): Ausgleichungsrechnung, statistische Auswertemethoden. 2. überarbeitete und erweiterte Auflage. de Gruyter, ISBN 978-3-11-019055-7
- Rao C R et al. (2008): Linear Models and Generalizations. Least Squares and Alternatives. 3rd extended edition. Springer, ISBN 978-3-540-74226-5

→ Contents

References

- Sneeuw N, F Krumm and M Roth (20xx): Lecture Notes Adjustment Theory. Institute of Geodesy, Stuttgart University
- Strang G (2006): Linear Algebra and its Applications. 4th edition. Thomson Learning, Inc., ISBN 0-03-010567-6
- Strang G (2009): Introduction to Linear Algebra. 4th edition. Wellesley-Cambridge Press, ISBN 078-0-9802327-1-4
- Strang G and K Borre (1997): Linear Algebra, Geodesy, and GPS. Wellesley-Cambridge Press, ISBN 0-9614088-6-3
- Teunissen P J G (2001): Dynamic data processing - recursive least-squares. Delft University Press, ISBN 13 978-90-407-1976-9
- Teunissen P J G (2003): Adjustment theory - an introduction. Delft University Press, ISBN 13 978-90-407-1974-5
- Teunissen P J G (2006a): Network quality control. Delft University Press, ISBN 13 978-90-71301-98-8
- Teunissen P J G (2006b): Testing theory - an introduction. Delft University Press, ISBN 13 978-90-407-1975-2
- Weber H (1992) Einführung in die Wahrscheinlichkeitsrechnung und Statistik für Ingenieure. Teubner, ISBN 3-519-02983-9
- Wolf H (1979) Ausgleichungsrechnung II. Aufgaben und Beispiele zur praktischen Anwendung. Dümmlers, ISBN 3-427-78361-8

→ Contents

Contents

- Introduction to adjustment theory: Physical model, mathematical model, redundancy, inconsistency, geodetic linear model, errors, historical review
- Least squares adjustment: Geometry, linear algebra, calculus, probability theory, basic approaches
- Approach I: A-model (parameter adjustment, adjustment of observation equations); geometry and linear algebra, calculus with examples; rank deficiency and the datum problem: parameter adjustment with constraints; linear regression with/without constraints
- Approach II: B-model (condition adjustment, adjustment of condition equations), geometry and linear algebra, calculus with examples
- Taylor series expansion of non-linear functions: Linearization of non-linear observation and condition equations, iterative adjustment
- Mixed model adjustment
- Weighted least squares
- Random variables, probability distributions
- Probabilistic approach A-model, error propagation
- Hypothesis testing, test distributions

→ Elaborated examples

Elaborated examples

- Buying goods in a super market
- Measuring a table length twice
- Height network observed by height differences (Leveling network, Fix/free datum)
- Planar triangle observed by angles
- Planar triangle observed by angles and distances
- Planar 4-point network observed by directions and distances
- Planar 9-point network observed by distances
- 3D-plane regression (A-model, mixed model)
- Various straight line fits (A-model, B-model, mixed model)
- Various planar coordinate transformations (A-model, mixed model)
- Various ellipse fits under different restrictions (Mixed model)
- Polynomial regression (with/without constraints)
- Non-linear function fit
- Error ellipses
- Testing apples to fit in trays

→ Introduction

Introduction

Adjustment theory deals with the optimal combination of redundant measurements/observations, together with the estimation of unknown parameters (Teunissen 2000)

Example: We go shopping twice 3 apples + 4 pears = 5 €
5 apples + 2 pears = 6 €

"Physical model" (Reality)

Unknowns: price per apple (x_1), price per pear (x_2)

Case 0:
$$3x_1 + 4x_2 = 5€$$
$$5x_1 + 2x_2 = 6€$$

"Mathematical model"

$$\begin{pmatrix} 3 & 4 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix} \quad \sim \quad Ax = y$$

"Matrix notation" "Linear Algebra"

$\det A = -14 \neq 0 \Rightarrow$ system is solvable and x_1, x_2 can be determined.

$$x = A^{-1}y$$
$$= -\frac{1}{14}\begin{pmatrix} 2 & -5 \\ -4 & 3 \end{pmatrix}' \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$
$$= \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}$$

Physical model: x_1=1 €/apple, x_2=50 cent/pear.

NE

→ Introduction

Introduction

Prices could be determined because number of observations (shopping twice) equals the number of unknowns (prices), and matrix A is regular (not singular) and thus invertible.

We have the linear model y=Ax

array of unknown parameters
matrix of coefficients (design matrix)
array of observations

<u>Case 1a:</u> Amount of apples and pears are a multiple the second shopping and one has to pay the same multiple ⇒ no new information is added

1st shopping:	3 apples + 4 pears = 5 €
2nd shopping:	6 apples + 8 pears = 10 €

$$\begin{pmatrix} 3 & 4 \\ 6 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 10 \end{pmatrix} \quad \sim \quad Ax = y$$

Matrix A is singular, rank deficient, has dependent rows (and columns): detA=0.
x cannot be determined.

→ Introduction

Introduction

<u>Case 1b:</u> Same configuration (A) as in case 1a but one has to pay $y=[5,8]'$. A is still not invertible but the price of the second shopping is only 8€, in addition. This is not consistent with the result we would have expected and not consistent with the other data; the price of the second shopping does not fit to the first shopping.

Though: This case of inconsistency is not necessarily a weakness.
May indicate mistakes in the overall bill.
May indicate an erroneous linear model: the prices may have changed in the meantime.

<u>Case 2:</u> Back to the invertible case but with a third <u>inconsistent</u> buying, y_3.

$$\begin{pmatrix} 5 \\ 6 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 5 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Combine 1.+2. shopping $\Rightarrow x_1=1, x_2=0.5$
(not consistent with 3rd shopping)

Combine 1.+3. shopping $\Rightarrow x_1=-1, x_2=2$
(not consistent with 2nd shopping)

Combine 2.+3. shopping $\Rightarrow x_1=3/4, x_2=9/8$
(not consistent with 1st shopping)

→ Introduction

Introduction

In the general case we have redundant and inconsistent information: the number of observations (equations = rows of A) > the number of unknowns (columns of A) $\Rightarrow$ A is a rectangular matrix and not invertible, the redundancy (number) is $r = m - n$.

But: Redundancy/inconsistency is a desirable feature in linear models ! It allows quality control of observations, strengthens the estimation of the unknowns, allows testing the unknown parameters.

Q: How to eliminate the inconsistencies ? How to solve such a system ?

A: Spread them out in an optimal way, i.e. combine redundant data optimally. $\rightarrow$ New question: What does 'optimally' mean ?

The inconsistencies in our examples were probably caused by a change in the prices between two rounds of shopping and could not be modelled in our model $y=Ax$; we should have introduced new parameters to account for the price change. In Geodesy, however, inconsistencies are mainly due to observation errors.

$\rightarrow$ Introduction

Introduction

In order to make a Geodetic Linear Model, we have to introduce an unknown vector e of inconsistencies (errors, residuals, discrepancies) transforming the inconsistent model into a consistent model.

$$\underset{m\times 1}{y} = \underset{m\times n}{A}\,\underset{n\times 1}{x} + \underset{m\times 1}{e} \quad , \quad e \neq 0$$

$$\left.\begin{matrix} y+e=Ax & \ell+v=Ax & y-e=X\beta \\ y+e=X\beta & E\{y\}=X\beta & \cdots \end{matrix}\right\} \text{ alternative representations}$$

When we talk about errors we should distinguish between

(i) gross errors (blunders, outliers)

(ii) systematic errors

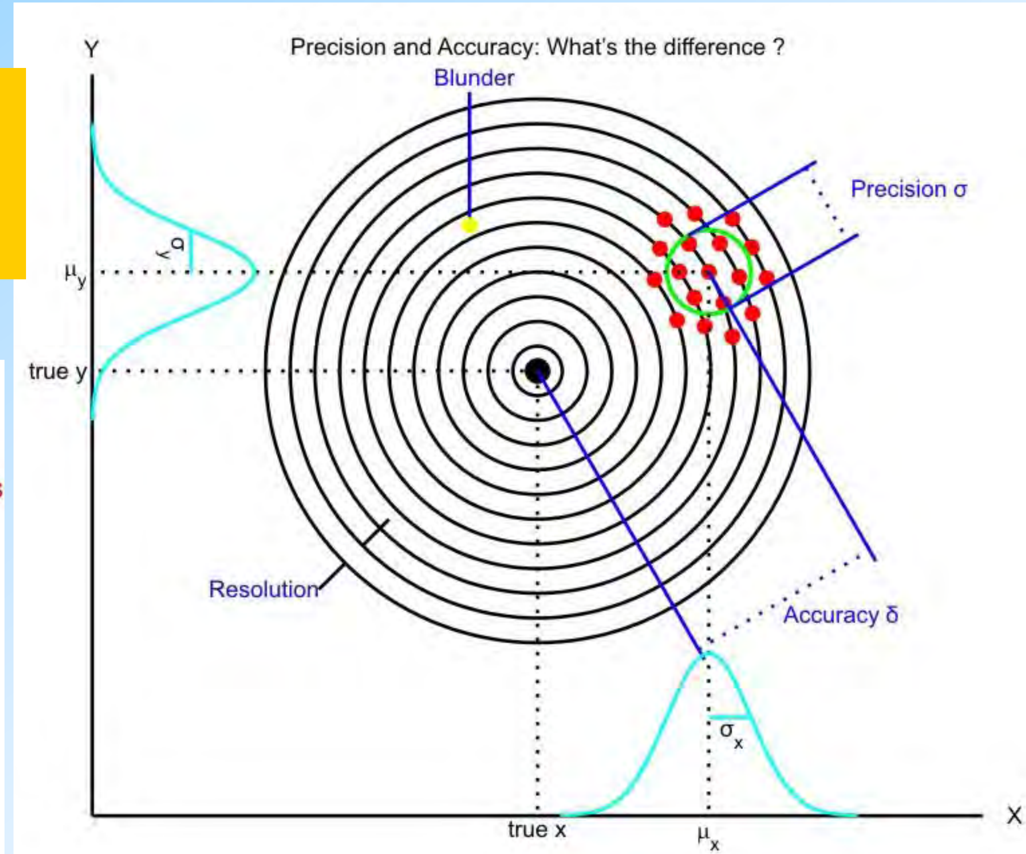
(iii) random errors

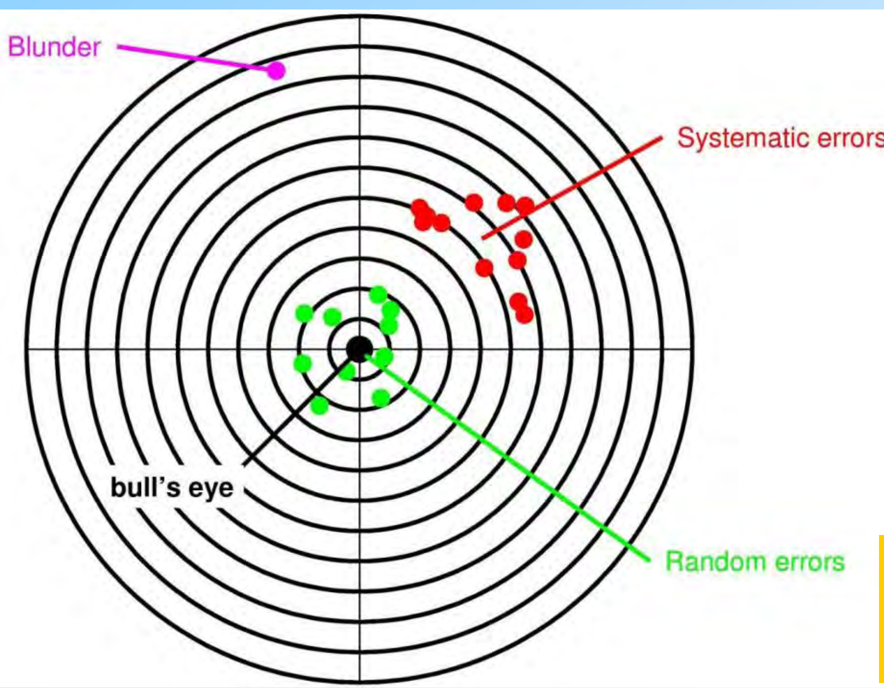
→ Introduction

Introduction

Marks left behind by an arrow in the game of darts, in which one tries to aim at bull's eye

Errors are always stochastic quantities → e is a stochastic variable → probability theory





irregular behaviour of marks around the center

→ Introduction

Introduction

Possible choices for e in the consistent model

$$\underset{m\times 1}{y} = \underset{m\times n}{A}\,\underset{n\times 1}{x} + \underset{m\times 1}{e} \quad \sim \quad \begin{pmatrix} 5 \\ 6 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 5 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

$$\begin{array}{llllll}
 & x_1=1 & , x_2=0.5 & \Rightarrow e_1=0 & , e_2=0 & , e_3=1 \\
 & x_1=0 & , x_2=1 & \Rightarrow e_1=1 & , e_2=4 & , e_3=1 \\
\text{Meaningless price} \rightarrow & x_1={\color{red}-1} & , x_2=2 & \Rightarrow e_1=0 & , e_2=7 & , e_3=0 \\
 & x_1=3/4 & , x_2=9/8 & \Rightarrow e_1=-7/4 & , e_2=0 & , e_3=0 \\
 & \ldots & & & & \\
 & \hat{x}_1=10/11 & , \hat{x}_2=89/132 & \Rightarrow \hat{e}_1=-28/66 & , \hat{e}_2=7/66 & , \hat{e}_3=49/66
\end{array}$$

What is special with the last solution ?

→ Historical review

Historical review

Observation model

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

Observations/measurements (Experiment outcome)

Linear relationship

linear combination matrix A (Design matrix)

Unknowns

Combine observations using a certain matrix L with as many rows as number of unknowns

$$\underset{n\times m}{L}\ \underset{m\times 1}{y} = \boxed{\underset{n\times m}{L}\ \underset{m\times n}{A}}\ \underset{n\times 1}{x} \qquad \operatorname{rank}(LA) = \operatorname{rk}(LA) = n$$

Solve for x

$$x = \underbrace{(LA)^{-1} L}\ y$$

$$\underset{n\times 1}{x} = \underset{n\times m}{C}\ \underset{m\times 1}{y}$$

Here, C is called the left inverse of A, because $CA=I_n$.
Depending on the choice of L, matrix C is not unique.

→ Historical review

Historical review

(ι) <1750: use only the first n out of m observations (method of selected points)

$$\underset{n\times m}{L} = \left[\underset{n\times n}{I} \quad \underset{n\times(m-n)}{0} \right]$$

Disadvantage: arbitrariness of the choice of n observations

Advantage: Cross-validate results using remaining observations

(ii) Combine observations in different other ways and average them

(iii) Use all combinations of n observations and look at the different solution vectors x, the number of which is $\binom{m}{n} = \frac{m!}{(m-n)!n!}$

(iv) Try to make the sum of all inconsistencies zero $\sum_{i=1}^{m} e_i = 0$ and $\sum_{i=1}^{m} |e_i| = \min$

(v) ~1800 (Legendre, Gauß): Minimize the square sum of residuals

$$\sum_{i=1}^{m} e_i^2 = \min \quad \Rightarrow \quad L = A'$$

$$\Downarrow$$

Least Squares Adjustment

→ Least squares adjustment

Least squares adjustment

Q: How can we combine observations y so that $\Omega := e'e = \sum_{i=1}^{m} e_i^2 = \min$ and why is L=A′ ?

A: Examine this question from different viewpoints

(i) Geometry: smallest distance $\Omega = e'e = (\underbrace{y - Ax}_{\text{distance}})'(y - Ax)$

(ii) Linear algebra: Orthogonality between (the columns of) A and e, A′e=0

(iii) Calculus: Minimizing e′e means differentiation of a target function

(iv) Probability theory: Best Linear Unbiased Estimate (BLUE)

→ Least squares adjustment

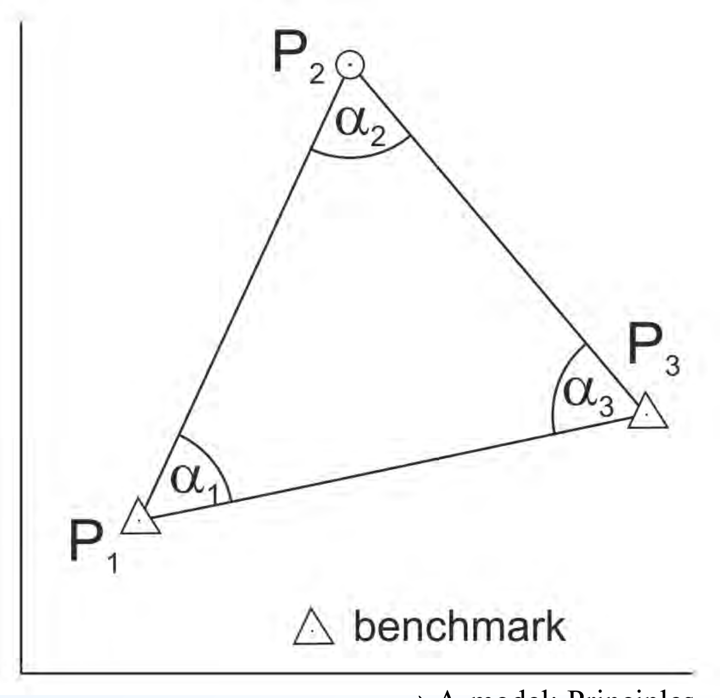
Least squares adjustment

Basic approaches to Least Squares Adjustment:

- Adjustment with observation equations (Parameter Adjustment, A-model, Gauß-Markoff model): Set up a (linear) relation between observations and unknown parameters (y=Ax+e). Example: Measure angles α_1, α_2, α_3 ("y") in order to determine coordinates ("x") of point P_2.
- Adjustment with condition equations (Condition Adjustment, B-model): No unknown parameters ("x") exist, use only observations ("y") to find e in order to 'correct' them (w=B′e). Example: Determine smallest possible corrections $e_{\alpha_1}, e_{\alpha_2}, e_{\alpha_3}$ so that

$$(\alpha_1 - e_{\alpha_1}) + (\alpha_2 - e_{\alpha_2}) + (\alpha_3 - e_{\alpha_3}) = 180° \,!$$

- Adjustment using a combination of both → Mixed model (Gauß-Helmert model)



→ A-model: Principles

A-model: Principles

<u>Example:</u> Determine the side length (x) of a table from measuring it twice (y_1,y_2) with a tape rule.

Observations $y_1 = x$ unknown table length

$y_2 = x$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} x$$

"direct observations"

$$\underset{m\times 1}{y} = \underset{m\times n}{A}\,\underset{n\times 1}{x}\,,\ \underset{2\times 1}{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix},\ \underset{m\times 1}{A} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

array notation

m: number of observations/measurements (here m=2)
n: number of unknown parameters ("unknowns") (here n=1)

<u>Discussion:</u>

- if $y_1=y_2$ (consistency) everything is fine and $x= y_1=y_2$
- if $y_1\neq y_2$ (inconsistency because of measurement errors) then m=2 unknown inconsistency parameters $e=[e_1,e_2]'$ must be added in order to make $y=Ax$ consistent: $y=Ax+e$

→ A-model: Principles

A-model: Principles

What is inconsistency ?

(i) The columns of A cannot be combined linearly (no x can be found) in order to obtain y. Example:

$$Ax = \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{a_1} x = xa_1 = \underbrace{x\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\text{Linear combination of column } a_1 \text{ of A}} \neq \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

(ii) The rank of A, q=rk(A) , is not equal to the rank of A augmented by y, rk([A|y])

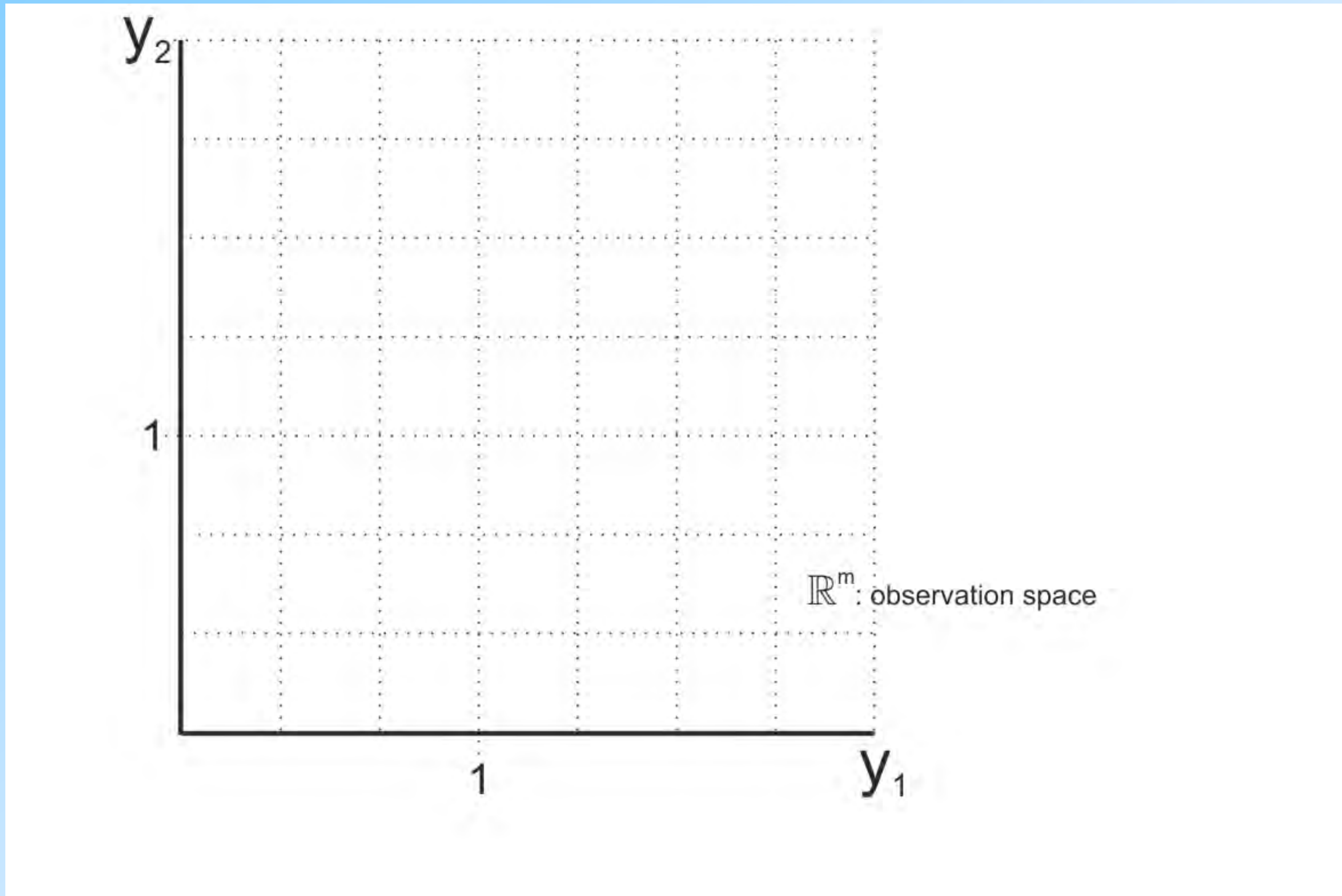
$$q = \mathrm{rk}(A) = 1, \quad \mathrm{rk}([A \mid y]) = \mathrm{rk}\begin{bmatrix} 1 & y_1 \\ 1 & y_2 \end{bmatrix} = 2 \quad \text{if } y_1 \neq y_2$$

(if y_1=y_2 then columns of A and y are linearly dependent ⇔ y is a multiple of the columns of A, rk([A|y])=1)

→ A-model: Geometry

A-model: Geometry

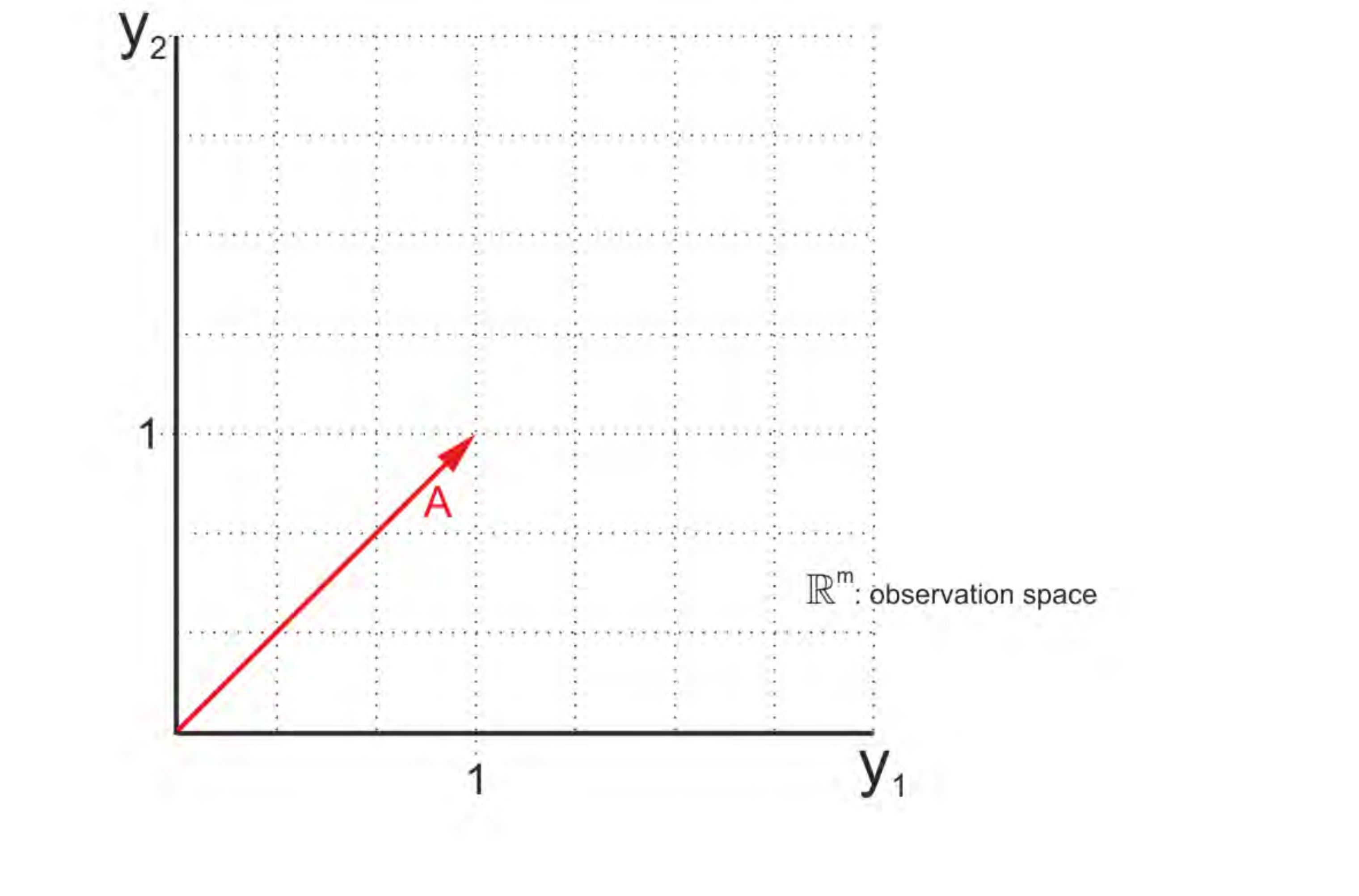
1) Create coordinate system with as many axes as observations, here m=2 → $\mathbb{R}^2$



→ A-model: Geometry

A-model: Geometry

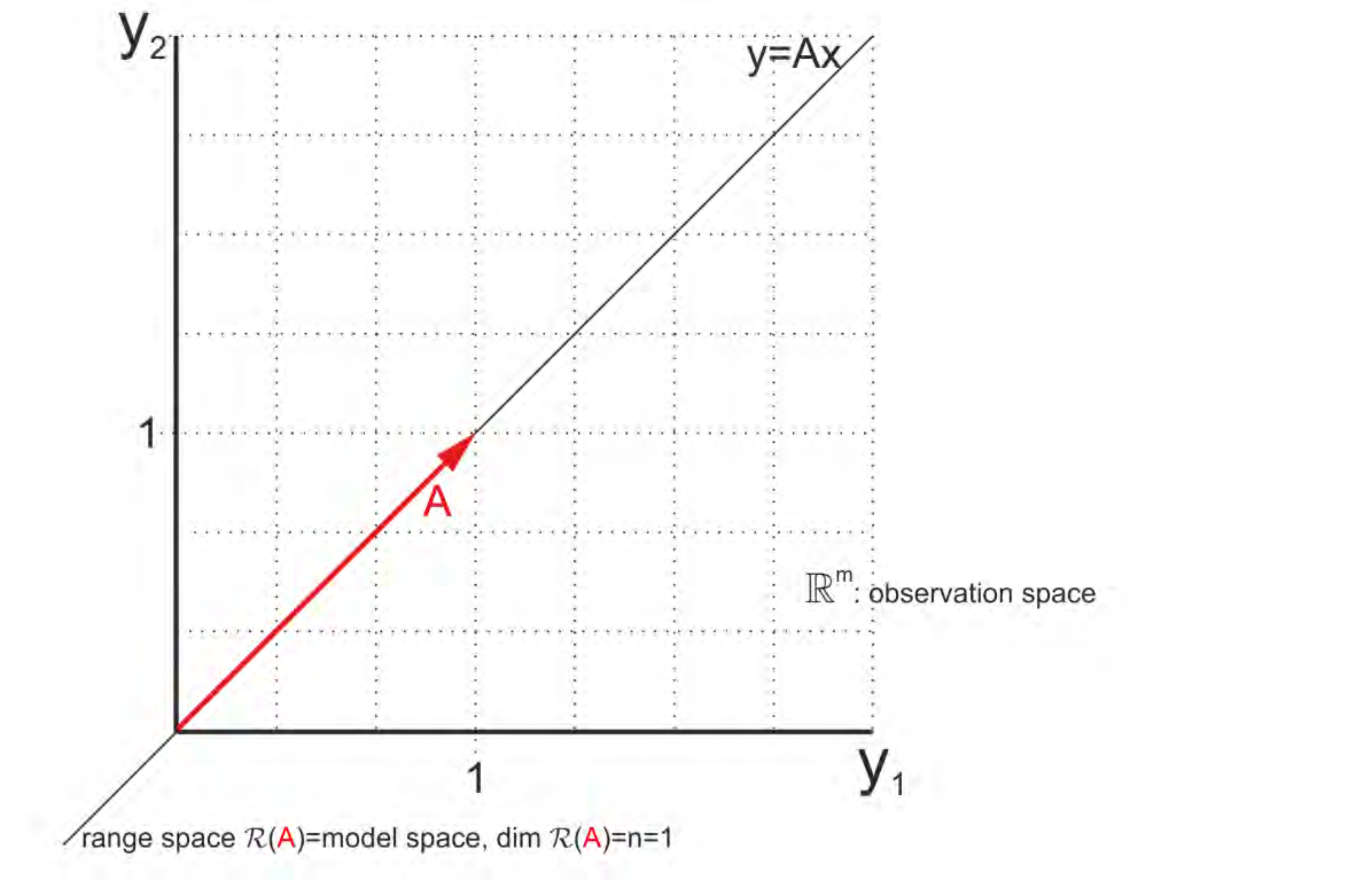
2) Plot columns of A, here A=[1,1]′ (a single column)



→ A-model: Geometry

A-model: Geometry

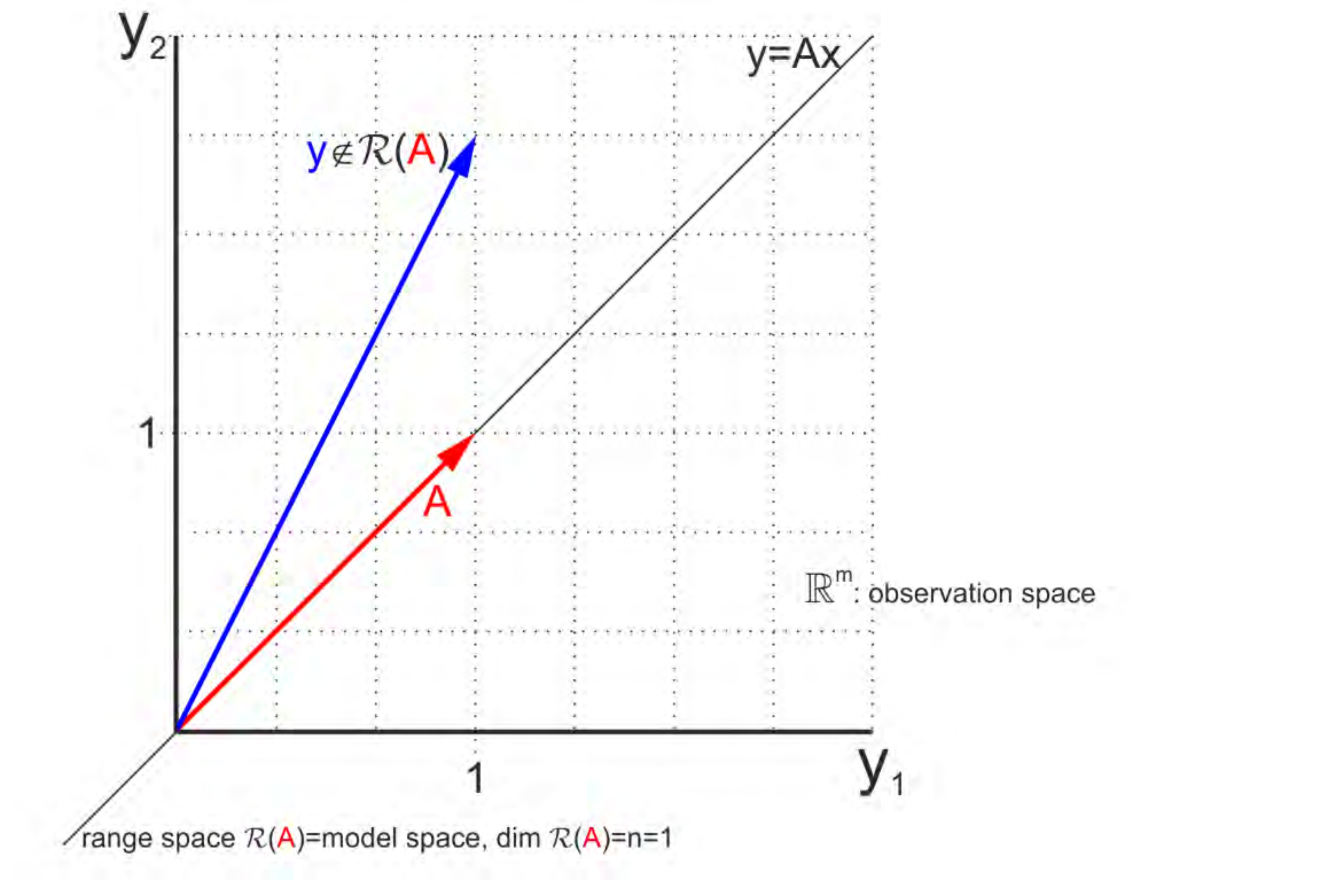
3) Draw y=Ax which is (here) a line in $\mathbb{R}^2$ spanned by A



→ A-model: Geometry

A-model: Geometry

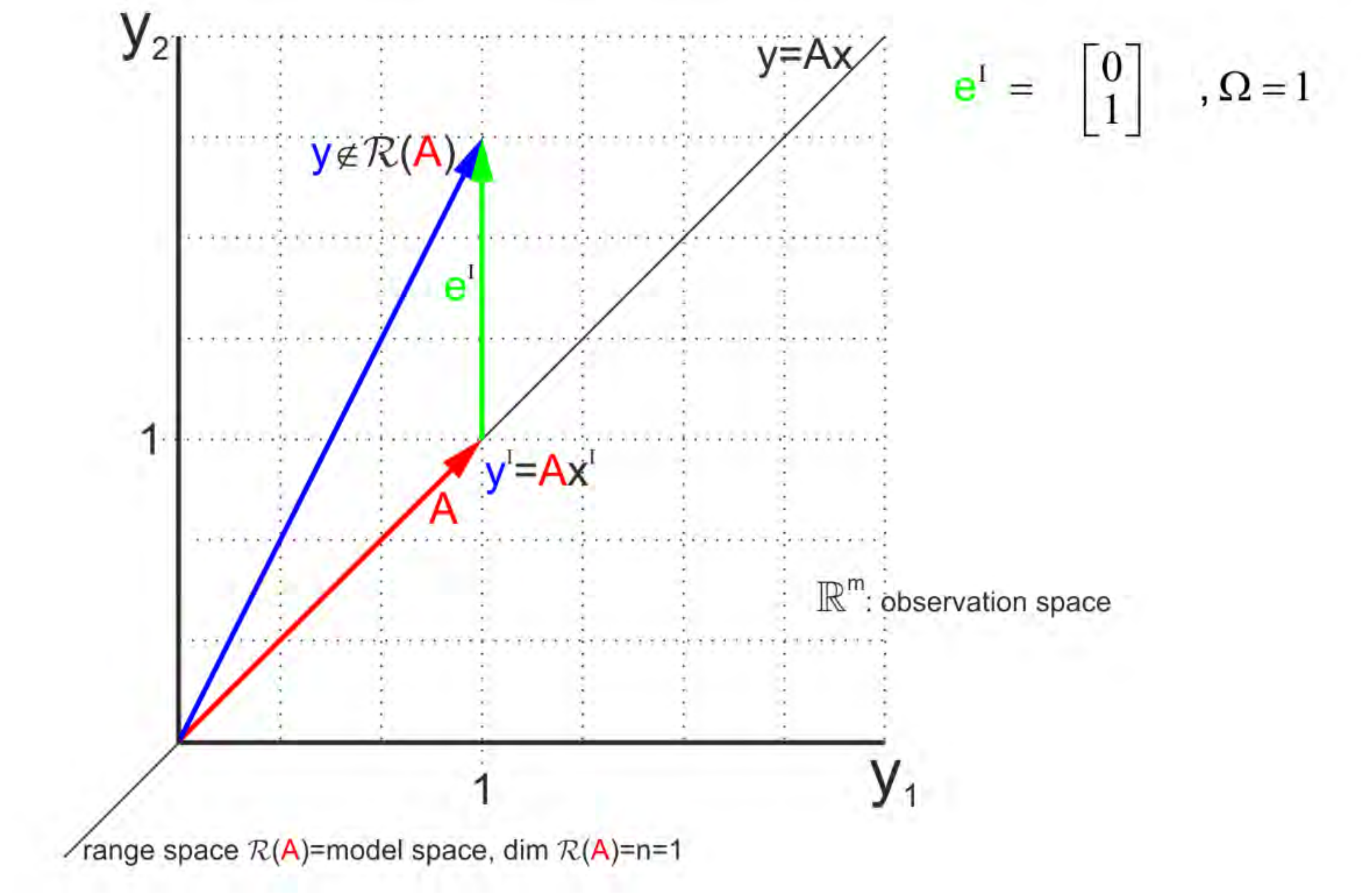
4) Plot (inconsistent vector of) observations y, which create a point in 2d-space



→ A-model: Geometry

A-model: Geometry

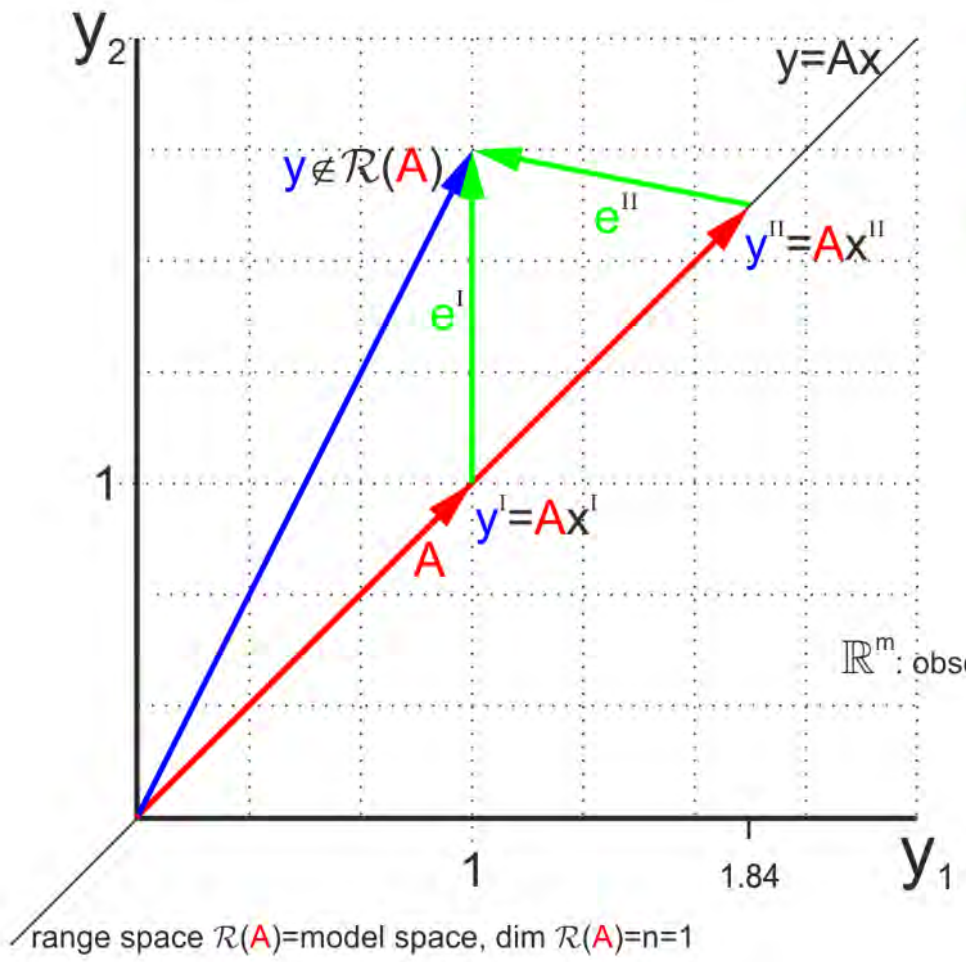
5) Find e's which satisfy y=Ax+e, corresponding x's and products Ax



→ A-model: Geometry

A-model: Geometry

5) Find e's which satisfy y=Ax+e, corresponding x's and products Ax



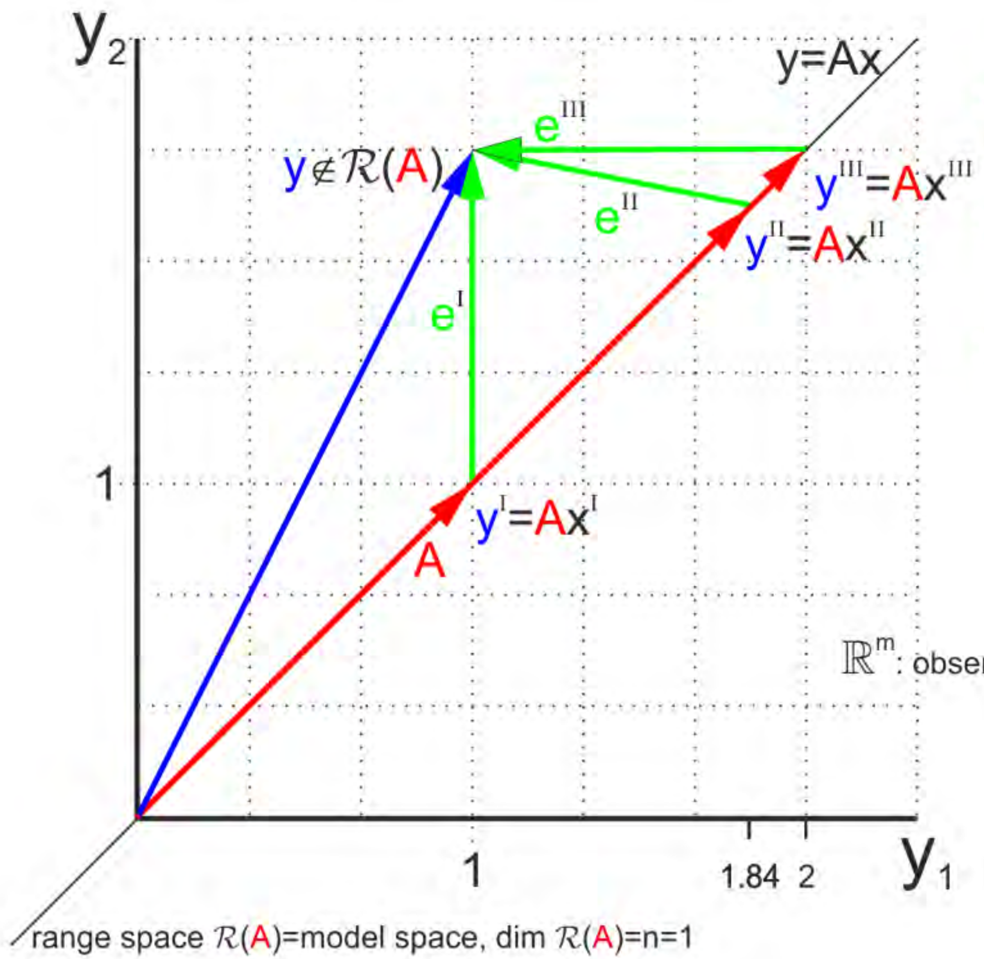
$$e^{I} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \Omega = 1$$

$$e^{II} = \begin{bmatrix} -0.84 \\ 0.16 \end{bmatrix}, \Omega = 0.73$$

→ A-model: Geometry

A-model: Geometry

5) Find e's which satisfy y=Ax+e, corresponding x's and products Ax



$$e^{I} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \Omega = 1$$

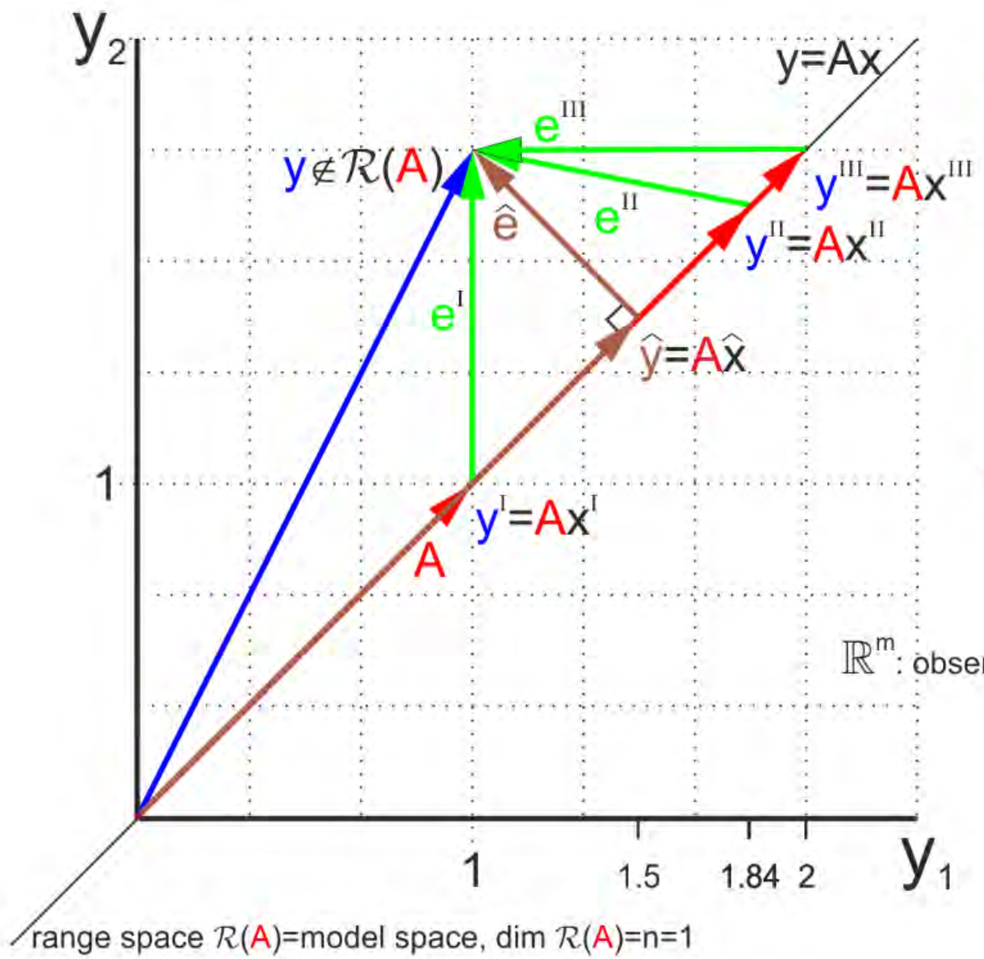
$$e^{II} = \begin{bmatrix} -0.84 \\ 0.16 \end{bmatrix}, \Omega = 0.73$$

$$e^{III} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \Omega = 1$$

→ A-model: Geometry

A-model: Geometry

5) Find e's which satisfy y=Ax+e, corresponding x's and products Ax



$$e^{I} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \Omega = 1$$

$$e^{II} = \begin{bmatrix} -0.84 \\ 0.16 \end{bmatrix}, \Omega = 0.73$$

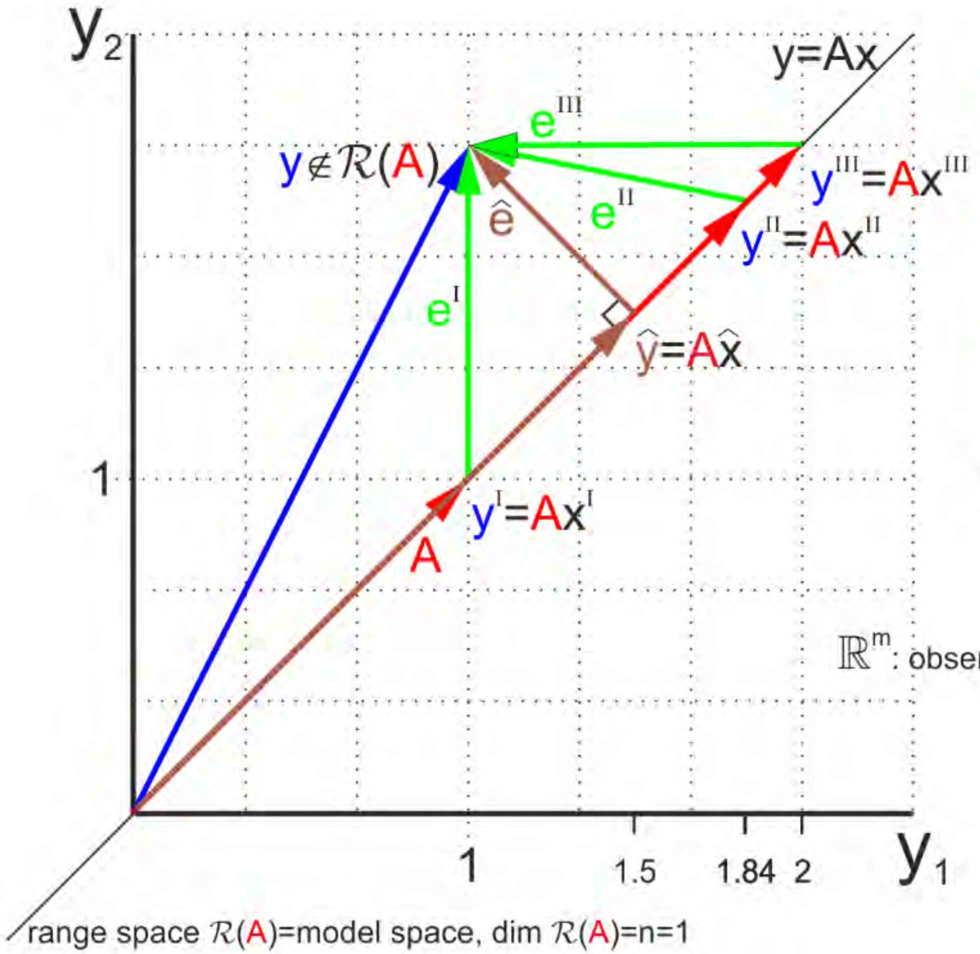
$$e^{III} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \Omega = 1$$

$$\hat{e} = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}, \Omega = 0.5$$

→ A-model: Geometry

A-model: Geometry

6) Choose shortest $\hat{e}$ which, in addition, satisfies least squares postulate $\hat{e}'\hat{e} = \min$



$$e^{I} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \Omega = 1$$

$$e^{II} = \begin{bmatrix} -0.84 \\ 0.16 \end{bmatrix}, \Omega = 0.73$$

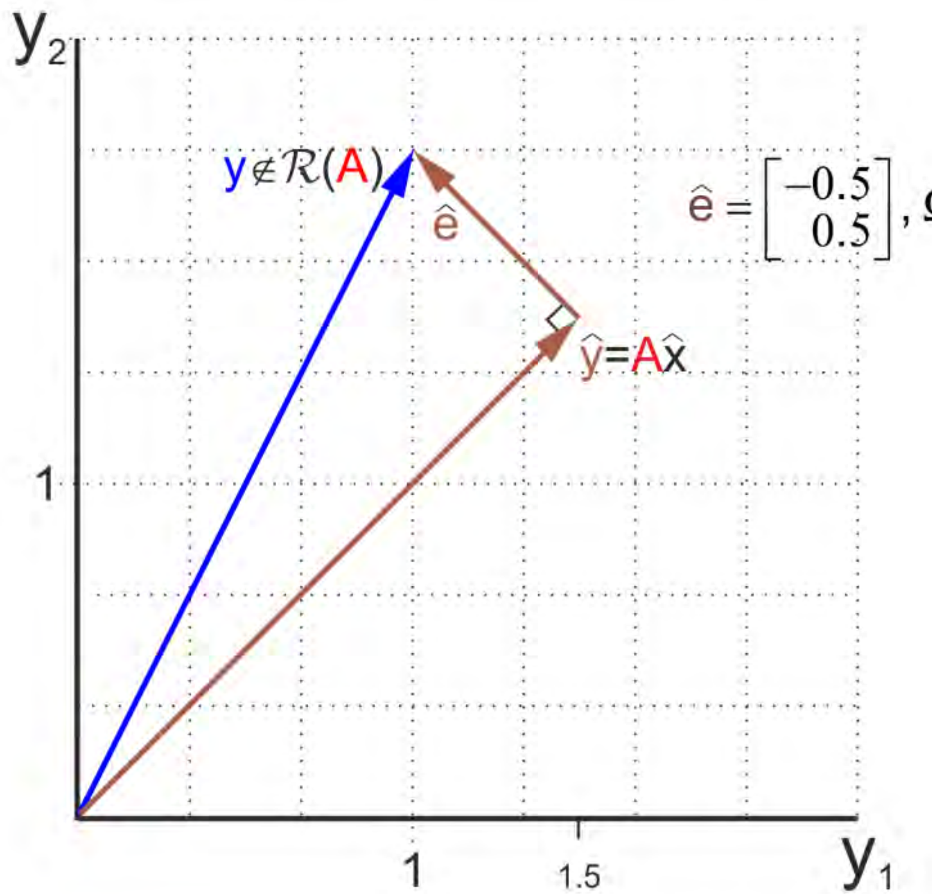
$$e^{III} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \Omega = 1$$

$$\hat{e} = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}, \Omega = 0.5$$

→ A-model: Geometry

A-model: Geometry

6) Choose shortest $\hat{e}$ which, in addition, satisfies least squares postulate $\hat{e}'\hat{e} = \min$



$$\hat{e} = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}, \Omega = 0.5 , \hat{x} = 1.84$$

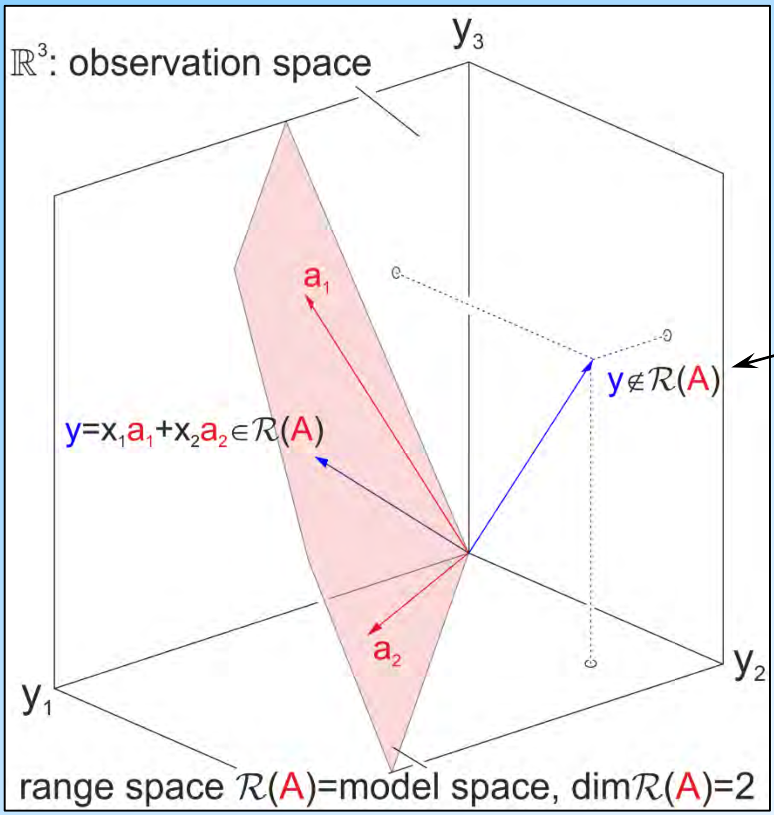
→ A-model: Geometry

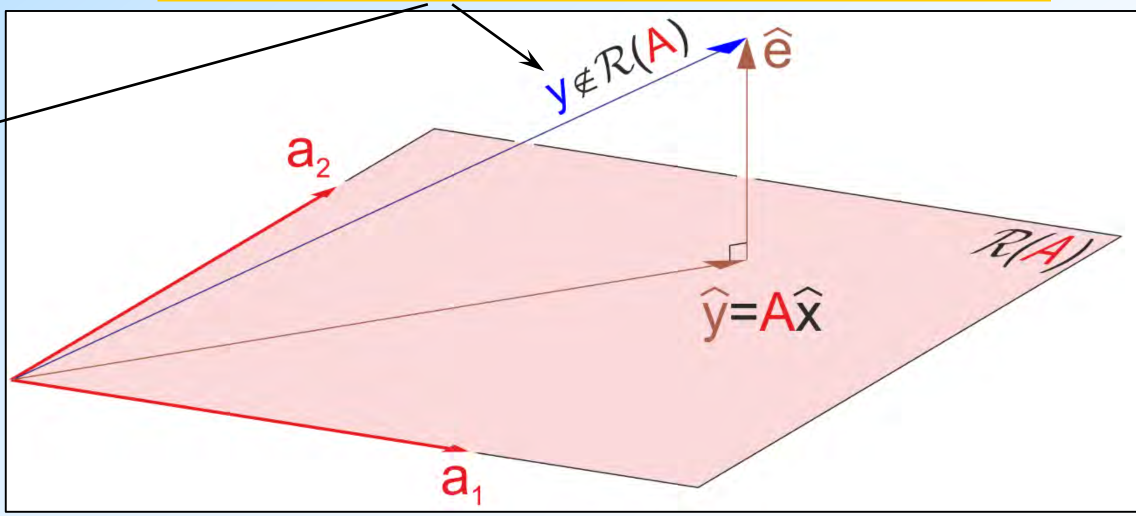
A-model: Geometry

Geometry in the case $m = 3$, $n = 2$: $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underset{3\times 2}{A} \underset{2\times 1}{x} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} = x_1 a_1 + x_2 a_2 = y$

$\begin{array}{cc} \uparrow & \uparrow \\ a_1 & a_2 \end{array}$

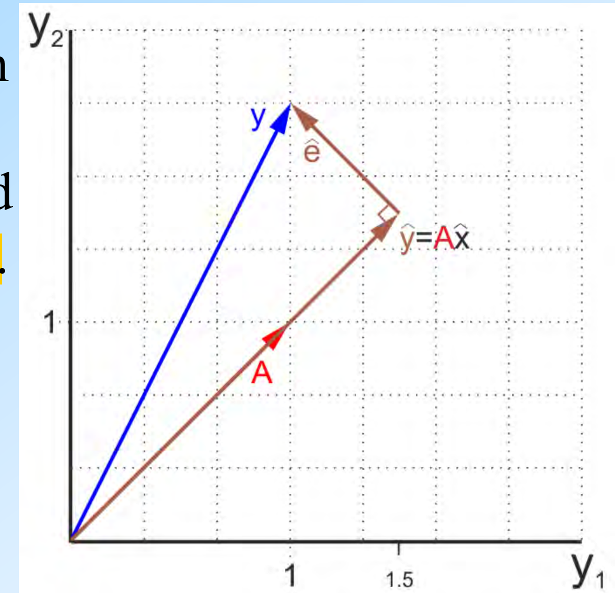
This vector y <u>is</u> in the range space of A because it is a linear combination of the columns a_1 and a_2 → consistency !
That vector y <u>is not</u> ! → Inconsistency !





→ A-model: Linear algebra

A-model: Linear algebra



For the shortest but still unknown $\hat{e}$ we have the equation $y=\hat{y}+\hat{e}=A\hat{x}+\hat{e}$. Left multiplication with A' yields $A'y=A'A\hat{x}+A'\hat{e}=A'A\hat{x}$ because A is orthogonal to $\hat{e}$ and the product $A'\hat{e}=0$. $A'A=N$ is the normal equation matrix.

$$
\begin{aligned}
&\Rightarrow A'\,y = A'A\hat{x} && \text{Normal equations}\\
&\Rightarrow \quad \hat{x} = \underbrace{(A'A)^{-1}A'}y = N^{-1}A'y && \text{LS-Estimate of } x\\
&\Rightarrow \quad \hat{y} = A\hat{x} && \text{LS-Estimate of } y\\
&\qquad\quad = A(A'A)^{-1}A'y && \text{Left inverse of } A\\
&\qquad\quad = P_A y && \text{Orthogonal projection of } y \text{ onto } A\\
&\Rightarrow \quad \hat{e} = y-\hat{y} && \text{LS-Estimate of } e \text{ (with } A'\hat{e}=0\text{)}\\
&\qquad\quad = y-A(A'A)^{-1}A'y\\
&\qquad\quad = \left[I_m - A(A'A)^{-1}A'\right]y\\
&\qquad\quad = (I_m - P_A)y = P_A^{\perp}y && \text{Orthogonal projection of } y \text{ onto a direction orthogonal to } A
\end{aligned}
$$

→ A-model: Linear algebra

A-model: Linear algebra

$\Rightarrow \hat{e}'\hat{e} = y'P_A^{\perp}y$ Square sum of estimated residuals/inconsistencies

$\Rightarrow$ Combination matrix $L = A'$

Projection matrices/projectors P_A and $P_A^{\perp}$ are idempotent matrices subject to

$$P_A = P_A' \;,\; P_A = P_A P_A \;,\; P_A^{\perp} = \left(P_A^{\perp}\right)' \;,\; P_A^{\perp} = P_A^{\perp} P_A^{\perp}$$

$$z = Iz = \left(P_A + P_A^{\perp}\right)z = P_A z + P_A^{\perp} z \quad \forall z \,,$$

i.e. any vector z can be decomposed into two orthogonal components, $P_A z$ and $P_A^{\perp} z$.

Proof:

$$P_A = A\left(A'A\right)^{-1}A' \Rightarrow P_A P_A = A\underbrace{\left(A'A\right)^{-1}A'A}_{\text{Identity matrix}}\left(A'A\right)^{-1}A' = A\left(A'A\right)^{-1}A' = P_A$$

$$P_A^{\perp} = (I - P_A) \Rightarrow P_A^{\perp}P_A^{\perp} = (I - P_A)(I - P_A) = I - P_A - P_A + \overbrace{P_A P_A}^{P_A} = I - P_A = P_A^{\perp}$$

→ A-model: Calculus

A-model: Calculus

Define the Lagrange- or cost function $\mathcal{L}_A(x) = 1/2 e'e$ and look for that $x = \hat{x}$ which minimizes $\mathcal{L}_A(x)$, i.e. $\mathcal{L}_A(\hat{x}) = \min$.

Because of the model equation y=Ax+e or y−Ax=e

$$\min_x \mathcal{L}_A(x) = \min_x \frac{1}{2} e'e = \frac{1}{2} \min_x e'e = \frac{1}{2} \min_x (y - Ax)'(y - Ax)$$

$$= \frac{1}{2} \min_x (y' - x'A')(y - Ax) = \frac{1}{2} \min_x (y'y - \underbrace{y'Ax}_{=x'A'y} - x'A'y + x'A'Ax)$$

$$= \frac{1}{2} \min_x (y'y - 2x'A'y + x'A'Ax)$$

Necessary condition for a minimum:

$$\frac{\partial \mathcal{L}_A(x)}{\partial x}(\hat{x}) = -A'y + A'A\hat{x} \overset{!}{=} 0 \quad \Rightarrow \quad \hat{x} = \left(A'A\right)^{-1} A'y \quad \dots \quad \Rightarrow \hat{y}, \hat{e} \quad \text{as before}$$

Sufficient condition for a minimum:

$$\frac{\partial^2 \mathcal{L}_A(x)}{\partial x^2}(\hat{x}) = A'A > 0 \quad \checkmark$$

→ A-model: Calculus

A-model: Calculus

Differentiation rules for vectors and matrices:
For the derivatives of a scalar-valued vector function f (a function with vectors/ matrices as arguments and returning a scalar) w.r.t. a vector the following general rules apply

$$f := a'x = x'a: \quad \frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(a'x) = \frac{\partial}{\partial x}(x'a) = a$$

$$f := x'Ax \quad : \quad \frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(x'Ax) = (A + A')x \overset{A=A'}{=} 2Ax$$

where both a and x are $k \times 1$ vectors and A is a $k \times k$ matrix.

For the Lagrange function $\mathcal{L}_A(x) = (y - Ax)'(y - Ax) = y'y - 2x'A'y + x'A'Ax$ the following differentiation results are achieved

$$f := y'y \Rightarrow \frac{\partial f}{\partial x} = 0$$

$$f := -2x'A'y \Rightarrow \frac{\partial f}{\partial x} = -2A'y$$

$$f := x'A'Ax = x'Nx \Rightarrow \frac{\partial f}{\partial x} = 2Nx = 2A'Ax$$

→ A-model: Example table length

A-model: Example table length

<u>Example:</u> Determine the side length (x) of a table from measuring it twice (y_1,y_2) with a tape rule.

Observations $y_1 = x$ unknown table length

$$y_2 = x$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} x$$

$$\underset{m\times1}{y} = \underset{m\times n}{A}\ \underset{n\times1}{x},\ \underset{2\times1}{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix},\ \underset{m\times1}{A} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \underset{1\times1}{A'A} = 2 \Rightarrow (A'A)^{-1} = 0.5,\ A'y = y_1 + y_2$$

$$\Rightarrow \hat{x} = \frac{1}{2}(y_1 + y_2)$$

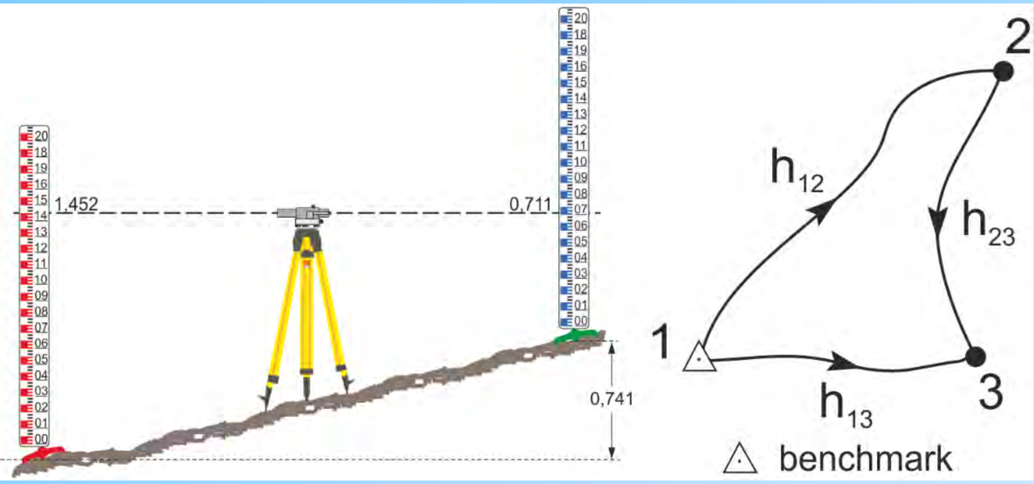
Mean average ("arithmetic mean") is a least squares estimate !

$$\Rightarrow \hat{y} = A\,\hat{x} = \frac{1}{2}\begin{bmatrix} y_1 + y_2 \\ y_1 + y_2 \end{bmatrix} \qquad \Rightarrow \hat{e} = y - \hat{y} = \frac{1}{2}\begin{bmatrix} y_1 - y_2 \\ -(y_1 - y_2) \end{bmatrix} \Rightarrow \hat{e}'\hat{e} = \frac{1}{2}(y_1 - y_2)^2$$

Check: $A'\hat{e} = 0$ ✓

→ A-model: Example height network

A-model: Example height network



Given: Benchmark with height H_1
Measured: Height differences h_{12}, h_{13}, h_{23}
Unknown(s): Heights H_2, H_3

Convention: $h_{ij} = H_j - H_i$

Do we have an adjustment problem ? Yes, because redundant – and therefore most probably inconsistent – information is available: m=3 measurements for n=2 unknown parameters H_2 and H_3.

<u>A-Model approach (m = 3 observations, n = 2 unknown parameters)</u>

Measurements →

Observation equations:

$$\left.\begin{aligned} h_{12} &= H_2 - H_1 \\ h_{13} &= H_3 - H_1 \\ h_{23} &= H_3 - H_2 \end{aligned}\right\} \Rightarrow \begin{bmatrix} h_{12} \\ h_{13} \\ h_{23} \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix}$$

→ A-model: Example height network

A-model: Example height network

The coefficient matrix $A = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}$ is called "design matrix". It reflects the appearance of the network. Now, as H_1 was assumed to be the height of a benchmark it can be subtracted on both sides of the equations leading to the so-called reduced observations equations with a new design matrix.

dot indicates inconsistency

$$\underset{3\times 1}{y} \doteq \begin{bmatrix} h_{12} + H_1 \\ h_{13} + H_1 \\ h_{23} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} H_2 \\ H_3 \end{bmatrix} = \underset{3\times 2}{A} \, \underset{2\times 1}{x}$$

Most probably, this system of equations will be inconsistent because of imperfect measurements h_{ij}.

no dot means consistency

$$\underset{3\times 1}{y} = \underset{3\times 2}{A} \, \underset{2\times 1}{x} + \underset{3\times 1}{e} \, , \quad \underset{3\times 1}{e} = \begin{bmatrix} e_{12} & e_{13} & e_{23} \end{bmatrix}'$$

A-model

The system of equations will be made consistent by adding an $m\times 1$ vector e of (unknown) inconsistencies, which "corrects" the observations h_{ij}. This is the formulation of the A-model.

→ A-model: Example height network

A-model: Example height network

Solution to the A-model: $\hat{x} = (A'A)^{-1}A'y, \quad \hat{y} = A\hat{x}, \quad \hat{e} = y - \hat{y}$

$$\Rightarrow \hat{x} = \begin{bmatrix} \hat{H}_2 \\ \hat{H}_3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 3H_1 + 2h_{12} - h_{23} + h_{13} \\ 3H_1 + 2h_{13} + h_{23} + h_{12} \end{bmatrix}, \; \hat{y} = \frac{1}{3}\begin{bmatrix} 3\hat{H}_2 \\ 3\hat{H}_3 \\ h_{13} - h_{12} + 2h_{23} \end{bmatrix}$$

LS-Estimates of x and y

$$\Rightarrow \hat{e} = \frac{1}{3}\begin{bmatrix} h_{12} - h_{13} + h_{23} \\ h_{13} - h_{12} - h_{23} \\ h_{12} - h_{13} + h_{23} \end{bmatrix}, \quad \begin{bmatrix} \hat{h}_{12} \\ \hat{h}_{13} \\ \hat{h}_{23} \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2h_{12} - h_{23} + h_{13} \\ h_{12} + 2h_{13} + h_{23} \\ h_{13} - h_{12} + 2h_{23} \end{bmatrix}$$

LS-Estimates of e and original observations

$$\Rightarrow A'\hat{e} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \checkmark$$

Orthogonality check

$$\Rightarrow \hat{y} - A\hat{x} = y - \hat{e} - A\hat{x} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \checkmark$$

Main check: "Corrected" (reduced) observations and estimated parameters must satisfy equations !

→ A-model: Example height network

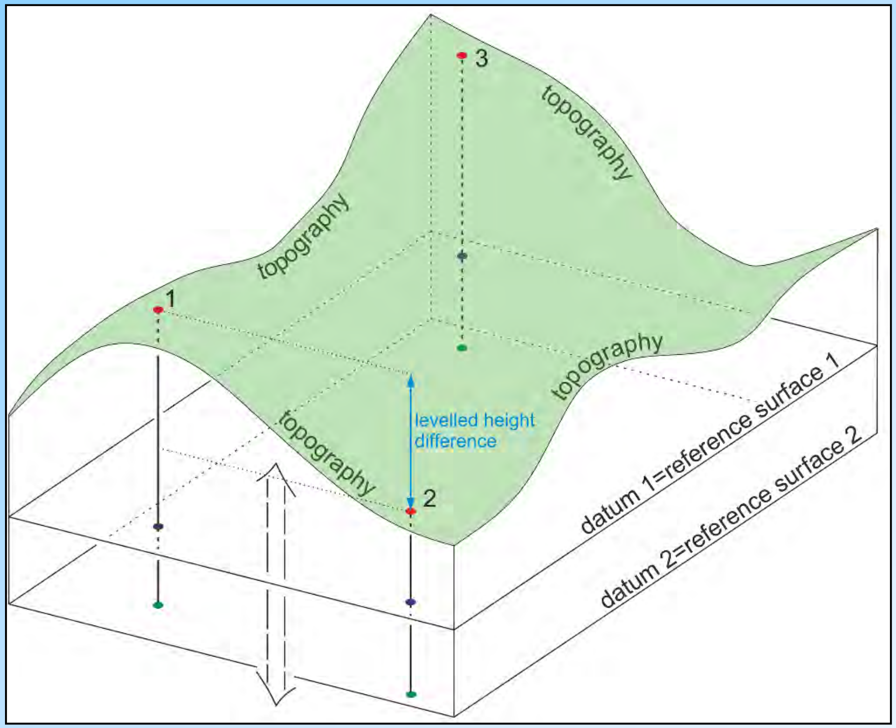
A-model: Example height network

Q: What happens in the A-model if H_1 was not given ?

A: H_1 could not be subtracted on both sides of the (original, i.e. non-reduced) observation equations.

As a consequence the original A-matrix reveals a rank deficiency: Col 1= –(Col 2+Col 3). The product N=A′A exhibits the same rank defect and will not be invertible. Height estimates and subsequent quantities cannot be computed. This is called a datum problem. It is caused by the fact that height difference observations (levelled height differences) are not sensitive to any constant shifts in height. They do not "feel" if the datum (the reference surface) is changed.

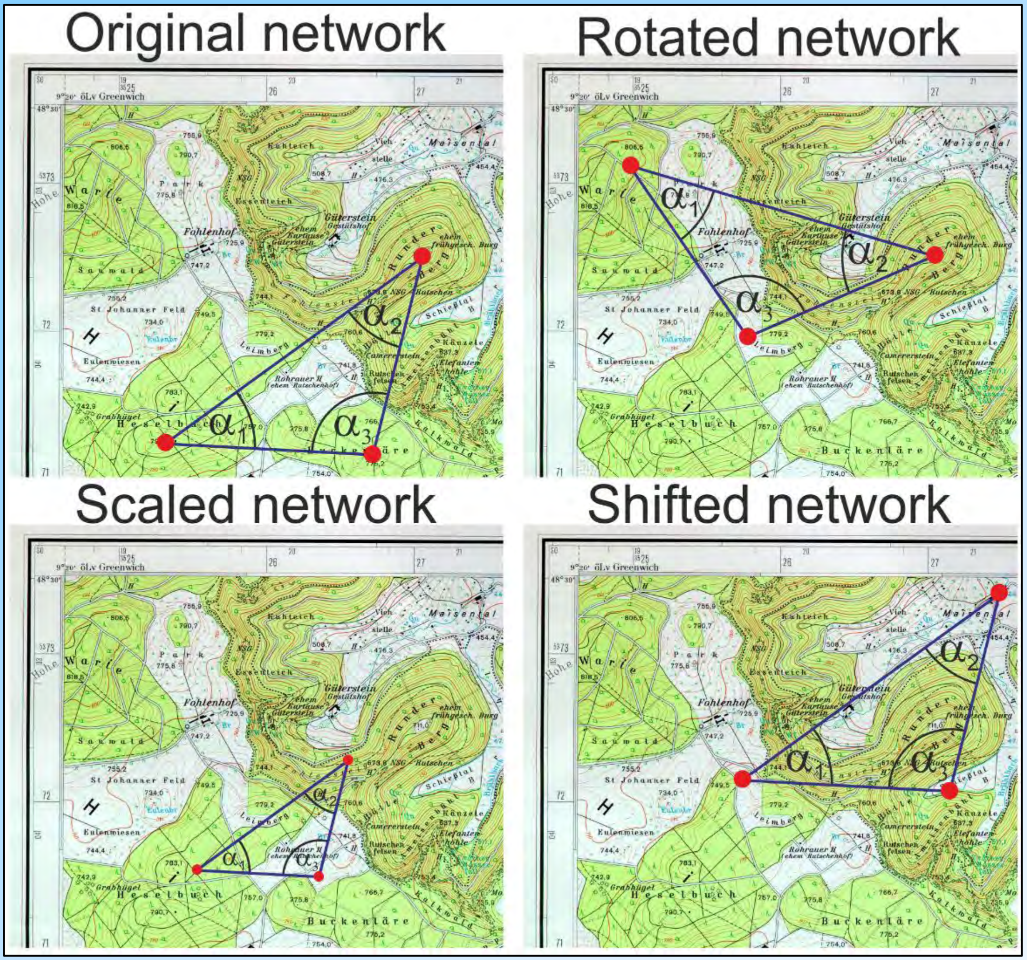
$$A = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}$$



→ A-model: Example trigonometric network

A-model: Example trigonometric network

In a 3-point trigonometric network (a planar triangle) the angles $\alpha_1,\alpha_2,\alpha_3$ have been measured with the purpose to determine the 2-d coordinates of the points.



Grundlage: Abbildungen – © Landesamt für Geoinformation und Landentwicklung Baden-Württemberg (www.lgl-bw.de) 06.07.2015, Az.: 2851.3-A/896

Q: Can the A-model do the job ?

A: No ! First m = 3 observations are not enough to determine n = 6 unknowns. Second – and this is even more serious – the datum is not properly defined: The network can be rotated, scaled and shifted without affecting the observations. Even in case of m > 6 measurements the m×n design matrix A has a rank deficiency of 4. Four coordinates would have to be fixed in order to remove the datum problem.

→ A-model: Datum problem

A-model: Datum problem

$$\underset{m\times 1}{y} = \underset{m\times n}{A}\ \underset{n\times 1}{x} + \underset{m\times 1}{e} \qquad m \geq n,\ s := \mathrm{rk}(A) \leq n$$

m: number of observations, n: number of unknowns, r = m–n: redundancy number, y: vector of (reduced) observations, A: design matrix, x: vector of unknown parameters, e: inconsistencies

Case 1: $\mathrm{rk}(A) = n:$ Target function $\mathcal{L}_A(x) = \frac{1}{2}e'e = \frac{1}{2}(y - Ax)'(y - Ax) \to \min_x$

$$\Rightarrow \hat{x} = N^{-1}A'y = (A'A)^{-1}A'y,\ \hat{y} = A\hat{x},\ \hat{e} = y - \hat{y}$$

N^{-1} exists because $\mathrm{rk}(A) = \mathrm{rk}(A'A) = \mathrm{rk}(N) = n$

Case 2 ("Datum problem"): $s := \mathrm{rk}(A) = \mathrm{rk}(A'A) = \mathrm{rk}(N) < n \Rightarrow N^{-1}$ does not exist ! The datum defect $d := n - s > 0$ is the dimension, dim $\mathcal{N}(A)$, of the nullspace $\mathcal{N}(A)$ of A. $\mathcal{N}(A)$ is the space spanned by the set of all solutions $x = x_{hom} \neq 0$ to the homogeneous equation $Ax=0$. The space $\mathcal{N}(A)$ contains more than only the zero vector if $s = \mathrm{rk}(A) < n$.

→ Excursion: Subspaces of a matrix A

Excursion: Subspaces of a matrix A

Let A be an m×n-matrix of arbitrary rank s: $\underset{m\times n}{A}$, $s := \text{rk}(A) \le \min(m, n)$. Then, there exist four fundamental subspaces called (1) column space $\mathcal{R}(A)$, (2) nullspace $\mathcal{N}(A)$, (3) row space $\mathcal{R}(A')$ and (4) left nullspace $\mathcal{N}(A')$. The row space of A equals the column space of A' and the left null space of A equals the null space of A'.

- $\mathcal{R}(A)$ is spanned by the s linearly independent columns of A (and all their linear combinations). Its dimension is $\dim\mathcal{R}(A) = s$. It is a subspace of $\mathbb{R}^m$ because the columns have m components.
- $\mathcal{R}(A')$ is spanned by the s linearly independent rows of A (and all their linear combinations). Its dimension is $\dim\mathcal{R}(A') = s$. It is a subspace of $\mathbb{R}^n$ because the rows have n components.

- $\mathcal{N}(A)$ is spanned by the n–s solutions x≠0 (and all their linear combinations) of the homogeneous equation Ax=0. Its dimension is $\dim\mathcal{N}(A) = n - s$. It is a subspace of $\mathbb{R}^n$ because all x's have n components.
- $\mathcal{N}(A')$ is spanned by the m–s solutions y≠0 (and all their linear combinations) of the homogeneous equation y′A=0 or A′y=0. Its dimension is $\dim\mathcal{N}(A') = m - s$. It is a subspace of $\mathbb{R}^m$ because all y's have m components.

→ Excursion: Subspaces of a matrix A

Excursion: Subspaces of a matrix A

- The row space is orthogonal to the nullspace, $\mathcal{R}(A') \perp \mathcal{N}(A)$, because the dot products of the rows of A with all x's are zero, Ax=0.
- The column space is orthogonal to the left nullspace, $\mathcal{R}(A) \perp \mathcal{N}(A')$, because the dot products of all y's with the columns of A vanish, y′A=0.
- The sum of the row space and the nullspace is $\mathcal{R}(A') \oplus \mathcal{N}(A) = \mathbb{R}^n$.
- The sum of the column space and the left nullspace is $\mathcal{R}(A) \oplus \mathcal{N}(A') = \mathbb{R}^m$.

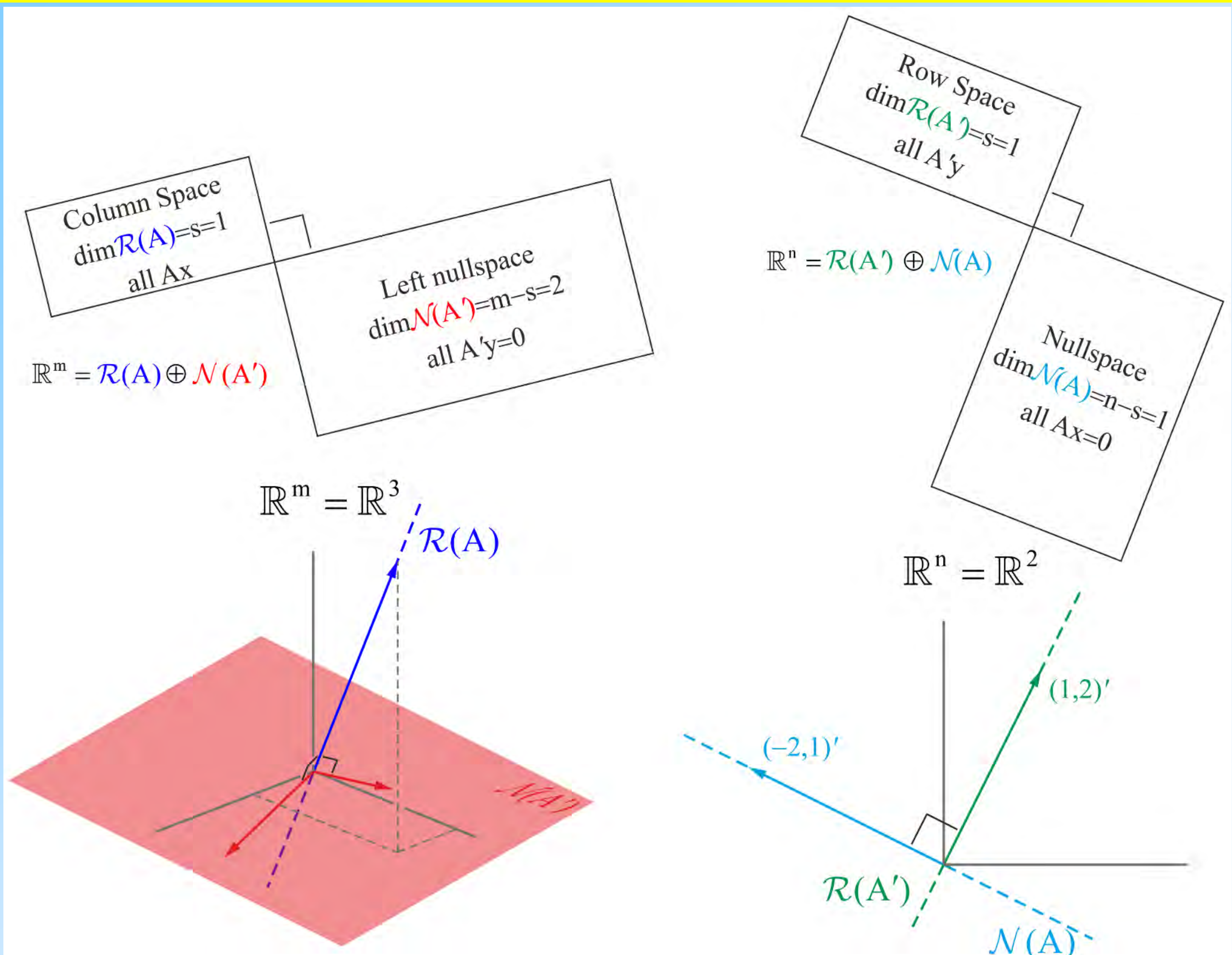
Example $(m = 3,\ n = 2,\ s = 1;\ \alpha, \beta, \gamma, \delta, \varepsilon \neq 0)$:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}, \qquad (A) = \alpha \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \qquad (A) = \beta \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \qquad (A') = \gamma \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\mathcal{N}(A') = \delta \begin{bmatrix} 1.185 \\ 0 \\ -0.395 \end{bmatrix} + \varepsilon \begin{bmatrix} -0.8 \\ 0.9 \\ -0.333 \end{bmatrix}$$

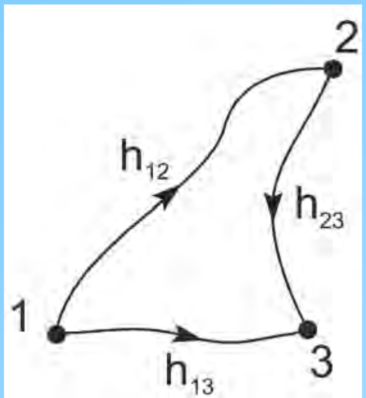
→ Excursion: Subspaces of a matrix A

Excursion: Subspaces of a matrix A



→ A-model: Datum problem

A-model: Example datum problem



In case of the 3-point height network without H_1 being fixed the system of equations $y=Ax+e$ cannot be solved uniquely for x because design matrix A is rank deficient. For any constant c, the vector $x_{hom}=c[1,1,1]'$ can be added and observations will not take notice: $y=A(x+x_{hom})+e=Ax+Ax_{hom}+e=Ax+e$. Vector $x_{hom}=c[1,1,1]'$ spans the nullspace $\mathcal{N}(A)$ of A.

$$A = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}$$

The same is true for the normal equation matrix $N=A'A=\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$, $x_{hom} = c\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$Nx_{hom}=A'Ax_{hom}=0$.

Remark: In a levelling network every diagonal element of N indicates how many levelling lines meet in that point/node. The off-diagonal numbers show if two points are connected by a levelling line. A zero means "no connection".

→ A-model: Datum problem

A-model: Datum problem

How to fix datum parameters using observables

(Translated from: Wolfgang Niemeier (2008): Ausgleichungsrechnung. 2. Auflage, de Gruyter, S. 233)

network Dimension	max. datum defect	observations	remaining datum defect d	remaining free parameters
1 Height Network	1	Height differences	1	Translation in z
		Absolute Heights	0	0
2 Planar Network	4	Directions+ Distances with one scale unknown	4	Translation in x and y Rotation around z Scale
		Distance(s) (+Directions)	3	Translation in x and y Rotation around z
		Azimuth(s) (+Directions)	3	Translation in x and y Scale
		Distances + Azimuths (+Directions)	2	Translation in x and y
		2 points with "absolute" coordinates	0	0

→ A-model: Datum problem

A-model: Datum problem

<table>
<tr><th>network Dimension</th><th>max. datum defect</th><th>observations</th><th>remaining datum defect d</th><th>remaining free parameters</th></tr>
<tr><td rowspan="6">3
3D-Network</td><td rowspan="6">7</td><td>Distances with one scale unknown
+Horizontal directions</td><td>7</td><td>Translation in x, y and z
Rotations around x, y and z
Scale</td></tr>
<tr><td>Distances+
Horizontal directions</td><td>6</td><td>Translation in x, y and z
Rotations around x, y and z</td></tr>
<tr><td>$\geq$ 2 Vertical directions* (with different orientation)
+ Horizontal directions</td><td>5</td><td>Translation in x, y and z
Rotation around z
Scale</td></tr>
<tr><td>Distances+
$\geq$ 2 Vertical directions*
+ Horizontal directions
+Azimuths</td><td rowspan="2">3</td><td rowspan="2">Translation in x, y and z</td></tr>
<tr><td>3 well scattered GPS-coordinate differences</td></tr>
<tr><td>3 well scattered points with "absolute" coordinates</td><td>0</td><td>0</td></tr>
</table>

*) $\equiv$ Zenith angles

→ A-model: Datum problem

A-model: Datum problem

Computationally, there are two alternative approaches to solve the datum problem:

Approach 1: Shrink the solution space by setting the height of <u>one</u> point to a constant, thus eliminating it as an unknown parameter from the vector x. One point is sufficient because d:=dim $\mathcal{N}(A)$= 1. By doing so the corresponding column is cancelled from A, thus causing the rank of A to become full. In the example before, H_1 was fixed and moved from the right side to the left side y. In general, it is eligible to fix more than one point but this will exert "pressure" on the network.

Approach 2: Augment the solution space by explicitly adding at least d constraints D′x=c to the system of linear equations. The solution is than found from the constrained Lagrangean

$$\text{Target function } \mathcal{L}_A(x,\lambda) = \frac{1}{2}\underset{1\times 1}{e'e} + \underset{1\times d}{\lambda'}(\underset{d\times n}{D'}\,\underset{n\times 1}{x} - \underset{d\times 1}{c}) =$$

$$= \frac{1}{2}(y-Ax)'(y-Ax) + \lambda'(D'x-c) \to \min_{x,\lambda}$$

→ A-model: Datum problem

A-model: Datum problem

D must be chosen in such a way that both A and D span the entire n-dimensional space, i.e. $\mathbb{R}^n = \quad (A') \oplus \quad (D) = \quad (A') \oplus \quad (A) \Leftrightarrow \mathrm{rk}([\underset{n\times m}{A'} \mid \underset{n\times d}{D}]) = n$. However, $D \perp A$, i.e. $\underset{m\times n}{A}\,\underset{n\times d}{D} = \underset{m\times d}{0}$ is not required for the solvability of the system.

Target function: $\mathcal{L}_A(x,\lambda) = \frac{1}{2}(y'y - 2y'Ax + x'A'Ax) + \lambda'(D'x - c) \rightarrow \min\limits_{x,\lambda}$

Necessary condition for a minimum:

$$\left.\begin{aligned}
\frac{\partial \mathcal{L}_A(x,\lambda)}{\partial x}(\hat{x},\hat{\lambda}) &= -A'y + A'A\hat{x} + D\hat{\lambda} = 0 \\
\frac{\partial \mathcal{L}_A(x,\lambda)}{\partial \lambda}(\hat{x},\hat{\lambda}) &= D'\hat{x} - c = 0
\end{aligned}\right\} \Rightarrow \underbrace{\begin{bmatrix} \underset{n\times n}{A'A} & \underset{n\times d}{D} \\ \underset{d\times n}{D'} & \underset{d\times d}{0} \end{bmatrix}}_{N^*} \begin{bmatrix} \underset{n\times 1}{\hat{x}} \\ \underset{d\times 1}{\hat{\lambda}} \end{bmatrix} = \underbrace{\begin{bmatrix} \underset{n\times n}{N} & \underset{n\times d}{D} \\ \underset{d\times n}{D'} & \underset{d\times d}{0} \end{bmatrix}}_{N^*} \begin{bmatrix} \underset{n\times 1}{\hat{x}} \\ \underset{d\times 1}{\hat{\lambda}} \end{bmatrix} = \begin{bmatrix} \underset{n\times 1}{A'y} \\ \underset{d\times 1}{c} \end{bmatrix}$$

While the normal equation matrix N is singular the extended normal equation matrix N* is regular and thus invertible. The system can be solved for $\hat{x}$ and all further quantities will follow as usual. In either case, the redundancy becomes r=m−(n−d).

Computational check: $\hat{e}'\hat{e} + y'A\hat{x} + c'\hat{\lambda} - y'y = 0$

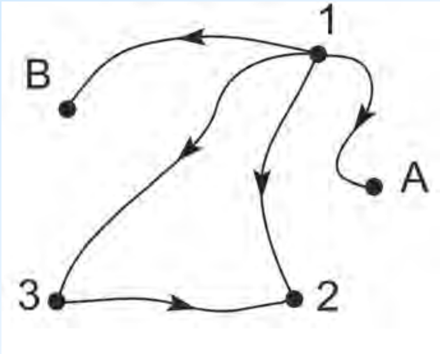
→ A-model: Datum problem

A-model: Datum problem

Remark 1: Approach 1 is disadvantageous as compared to approach 2 because design matrix A has to be changed whenever the datum definition is chosen differently.

Remark 2: More general constraints can only be incorporated in approach 2, see examples.

Example: "5-point height network" with $x = [H_A \; H_1 \; H_2 \; H_3 \; H_B]'$



(a) $D' = [0 \;\; 0 \;\; 0 \;\; 0 \;\; 1],\ c = 100$ removes the datum problem. H_B is used as a benchmark with height c=100.

(b) $D' = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix},\ c = \begin{bmatrix} 20 \\ 10 \end{bmatrix}$ overconstrains the solution because more than the minimum number d = 1 of datum constraints are specified: c(1)=20 is the benchmark height H_B and c(2)=10 is the benchmark height H_A.

(c) $D' = [1 \; 0 \; 0 \; 0 \; -1],\ c = 50$ puts $H_A = H_B + 50$. The datum problem is not removed, the rank test fails, $\mathrm{rk}([A' \mid D]) = 4 < 5$. The reference surface can still be shifted.

(d) $D' = [1 \;\; 1 \;\; 1 \;\; 1 \;\; 1],\ c = 10$ sets the barycentric height to 2. This D' is also used for a so-called free datum approach with no exterior forces.

→ A-model: Free datum

A-model: Free datum

A free adjustment (or a free datum choice) is based on the idea to let no external fixing act on the network/observations. The adjustment is solely defined by observations and approximate point coordinates/heights. Geometrically, the network is mapped onto the set of approximate coordinates/heights in such a way that the sum of squares of coordinate/height corrections becomes a minimum. The term MINOLESS (minimum-norm least squares solution) is sometimes used in this context. Geometrically, this can be seen as differential Helmert transformation.

While adjusted (reduced) observations, estimated inconsistencies and estimated square sum of residuals are invariant with respect to the choice of approximate coordinates/heights, adjusted point coordinates/heights strongly depend on it.

The overall statistical uncertainty of the adjusted coordinates/heights – expressed in terms of their variances – is smallest among all other solutions. The solution of a free network adjustment can be achieved using various algebraic approaches and algorithms (A-model with constraints, pseudoinverses, generalized inverses, rank partitioning's, approaches using a basis of the nullspace of A or N), which all give identical numerical results !

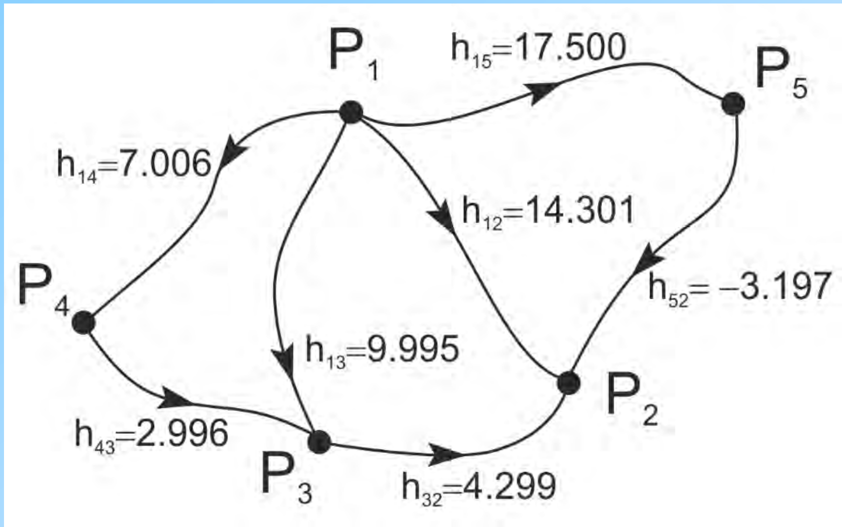
→ Free datum: Example Height Network

Free datum: Example Height Network

Approximate Heights H^0 [m]*

$H_1^0 = 93.459$
$H_2^0 = 107.759$
$H_3^0 = 103.459$
$H_4^0 = 100.459$
$H_5^0 = 110.956$
$\sum H_i^0 = 516.092$

* Least squares solution $\hat{H}$ will depend on H^0



Free height network adjustment

Measurements and reduced observations

$i-j$	h_{ij} [m]	$h_{ij} - (H_j^0 - H_i^0)$ [m]
$1-2$	14.301	0.001
$1-3$	9.995	–0.005
$1-4$	7.006	0.006
$1-5$	17.500	0.003
$3-2$	4.299	–0.001
$4-3$	2.996	–0.004
$5-2$	–3.197	0.000

Observation equation $\quad h_{ij} = H_j - H_i = H_j^0 + \Delta H_j - (H_i^0 + \Delta H_i)$

Reduced observation equation $\quad y := h_{ij} - (H_j^0 - H_i^0) = \Delta H_j - \Delta H_i$

Unknowns $\quad x = [\Delta H_1, \Delta H_2, \ldots, \Delta H_5]'$

Free network postulate $\quad D'x = \Delta H_1 + \ldots + \Delta H_5 = 0 \Leftrightarrow D' = [1 \quad 1 \quad 1 \quad 1 \quad 1]$

→ Free datum: Example Height Network

Free datum: Example Height Network

$$A = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 \end{bmatrix},\ N = \begin{bmatrix} 4 & -1 & -1 & -1 & -1 \\ -1 & 3 & -1 & 0 & -1 \\ -1 & -1 & 3 & -1 & 0 \\ -1 & 0 & -1 & 2 & 0 \\ -1 & -1 & 0 & 0 & 2 \end{bmatrix},\ A'y = \begin{bmatrix} -0.005 \\ 0.000 \\ -0.008 \\ 0.010 \\ 0.003 \end{bmatrix}_{[\mathrm{m}]},\ \begin{bmatrix} N & D \\ D' & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} A'y \\ 0 \end{bmatrix}$$

	$\hat{x}$ [mm]	$\hat{H}$ [m]
1	−1.000	93.458 0
2	−0.857	107.758 1
3	−2.143	103.456 9
4	3.429	100.462 4
5	0.571	110.956 6
	$\sum \hat{x}_i = 0$	$\sum \hat{H}_i = 516.092$

$i-j$	$\hat{h}_{ij}$ [m]*	$\hat{e}_{ij}$ [mm]*
1 − 2	14.300 14	0.857
1 − 3	9.998 96	−3.857
1 − 4	7.004 43	1.571
1 − 5	17.498 57	1.429
3 − 2	4.301 29	−2.286
4 − 3	2.994 43	1.571
5 − 2	− 3.198 43	1.429

* Invariant with respect to the choice of H^0

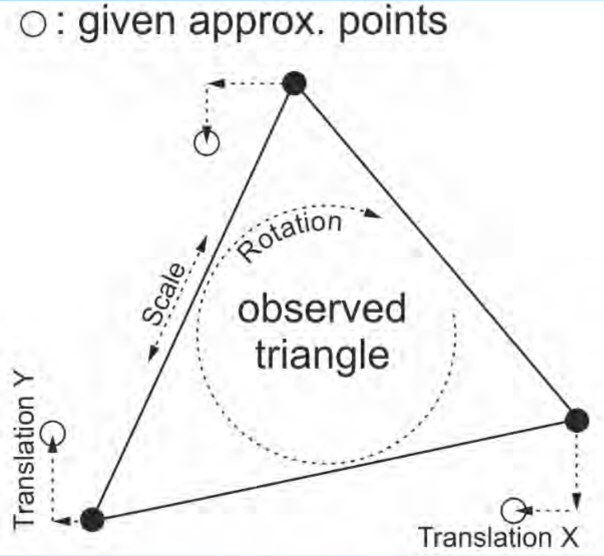
Square sum of residuals*
$\hat{e}'\hat{e} = 29.857\ \mathrm{mm}^2$

Orthogonality check
$A'\hat{e} = 0$ ✓

→ A-model: Free datum

A-model: Free datum

In case of a 2D network, which has been observed by angles/directions only, four constraints have to be specified in order to remove the rank deficiency of A and N. For a free adjustment these can be derived from a differential Helmert transformation of the set of (unknown) adjusted points onto the set of (given) approximate points. Again, adjusted point coordinates $\hat{x}$, $\hat{y}$ depend on the choice of x^0, y^0 while adjusted observations, estimated inconsistencies and square sum of residuals are independent.



Parameters of a differential Helmert transformation: Translation, rotation and scale

$$\sum(X_i - X_i^0) = \sum \Delta X_i = 0 \text{ "Translation in X"} \quad , \sum(Y_i - Y_i^0) = \sum \Delta Y_i = 0 \text{ "Translation in Y"}$$

$$\sum(Y_i^0 \Delta X_i - X_i^0 \Delta Y_i) = 0 \text{ "Rotation around Z"} \; , \sum(X_i^0 \Delta X_i + Y_i^0 \Delta Y_i) = 0 \text{ "Scale"}$$

For a scenario with measured distances (without scale unknown) the "scale constraint" will be disregarded. If the rotation of the network is fixed by an azimuth observation the "rotation equation" will be omitted.

→ A-model: Free datum

A-model: Free datum

Network type	# unknowns per point	d=dim $\mathcal{N}(A)$	Observation type(s)	$\mathcal{N}(A)$, type of datum defect		
				Translation	Rotation	Scale
Levelling	1 (height)	1	Height differences	1		
2D planar	2 (coordinates)	2	Distances and azimuths	$\begin{matrix}1 & 0\\ 0 & 1\end{matrix}$		
		3	Distances	$\begin{matrix}1 & 0\\ 0 & 1\end{matrix}$	$\begin{matrix}Y_i^0\\ -X_i^0\end{matrix}$	
		4	Angles and/or distance ratios	$\begin{matrix}1 & 0\\ 0 & 1\end{matrix}$	$\begin{matrix}Y_i^0\\ -X_i^0\end{matrix}$	$\begin{matrix}X_i^0\\ Y_i^0\end{matrix}$
3D spatial	3 (coordinates)	3	Distances, azimuths, astronomical latitudes and longitudes	$\begin{matrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{matrix}$		
		6	Distances	$\begin{matrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{matrix}$	$\begin{matrix}0 & -Z_i^0 & Y_i^0\\ Z_i^0 & 0 & -X_i^0\\ -Y_i^0 & X_i^0 & 0\end{matrix}$	
		7	Angles and/or distance ratios	$\begin{matrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{matrix}$	$\begin{matrix}0 & -Z_i^0 & Y_i^0\\ Z_i^0 & 0 & -X_i^0\\ -Y_i^0 & X_i^0 & 0\end{matrix}$	$\begin{matrix}X_i^0\\ Y_i^0\\ Z_i^0\end{matrix}$

→ A-model: Free datum

A-model: Free datum

Inconsistent observation equations with datum defect:

$$\underset{m\times 1}{y} = \underset{m\times n}{A}\,\underset{n\times 1}{x} + \underset{m\times 1}{e} \qquad m \geq n,\ s := \mathrm{rk}(A) < n \qquad (d = n - s > 0)$$

Solution approach 1): Augment the solution space using special constraints

Example 1: 2D-network (p points, n=2p unknowns) observed by angles & distances, datum defect d=3 (ordering of unknowns: $\Delta X_1, \Delta Y_1, \ldots, \Delta X_p, \Delta Y_p$)

$$\underset{d\times n}{D'} = \begin{bmatrix} 1 & 0 & 1 & 0 & \cdots & 1 & 0 \\ 0 & 1 & 0 & 1 & \cdots & 0 & 1 \\ \hline Y_1^0 & -X_1^0 & Y_2^0 & -X_2^0 & \cdots & Y_p^0 & -X_p^0 \end{bmatrix}, \quad \underset{d\times 1}{c} = \begin{bmatrix} 0 \\ 0 \\ \hline 0 \end{bmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \end{matrix}\right\} \text{Translations} \\ \text{Rotation} \end{matrix}$$

Example 2: 2D-network as before but observed by angles only, datum defect d=4)

$$\underset{d\times n}{D'} = \begin{bmatrix} 1 & 0 & 1 & 0 & \cdots & 1 & 0 \\ 0 & 1 & 0 & 1 & \cdots & 0 & 1 \\ \hline Y_1^0 & -X_1^0 & Y_2^0 & -X_2^0 & \cdots & Y_p^0 & -X_p^0 \\ \hline X_1^0 & Y_1^0 & X_2^0 & Y_2^0 & \cdots & X_p^0 & Y_p^0 \end{bmatrix}, \quad \underset{d\times 1}{c} = \begin{bmatrix} 0 \\ 0 \\ \hline 0 \\ \hline 0 \end{bmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \end{matrix}\right\} \text{Translations} \\ \text{Rotation} \\ \text{Scale} \end{matrix}$$

→ A-model: Free datum

A-model: Free datum

Solution approach 2): Rank factorization of design matrix A

Any matrix that is not of full rank can be expressed as the product of a matrix F of full column rank and a matrix G of full row rank, i.e. the columns of F / the rows of G are linearly independent: $\underset{m\times n}{A} = \underset{m\times s}{F}\,\underset{s\times n}{G}\,, \quad s := \mathrm{rk}(A)$

Then the minimum-norm least squares solution (MINOLESS) can be solved in two steps.

Step 1: Minimize e'e and solve for z in the model equation

$$\underset{m\times 1}{y} = \underset{m\times n}{A}\,\underset{n\times 1}{x} + \underset{m\times 1}{e} = \underset{m\times s}{F}\,\underset{s\times n}{G}\,\underset{n\times 1}{x} + \underset{m\times 1}{e} = \underset{m\times s}{F}\,\underset{s\times 1}{z} + \underset{m\times 1}{e} \quad , \quad \underset{s\times 1}{z} := \underset{s\times n}{G}\,\underset{n\times 1}{x}$$

$$A = FG \Rightarrow F'A = (\underset{\substack{s\times m\ m\times s\\ \text{full rank}\\ \text{matrix}}}{F'\ F})G \Rightarrow G = \underbrace{(\underset{s\times m}{F'}\ \underset{m\times s}{F})^{-1}\ \underset{s\times m}{F'}}_{\text{"left inverse of F"}} A$$

$$\Downarrow$$

$$\hat{z} = (F'F)^{-1}F'y \quad , \quad \hat{y} = F\hat{z} \quad , \quad \hat{e} = y - \hat{y}$$

→ A-model: Free datum

A-model: Free datum

Step 2: Minimize $x'x$ under the constraint $Gx = (F'F)^{-1}F'y \sim Gx - (F'F)^{-1}F'y = 0$ and solve for x.

$$\mathcal{L}(x,\lambda) = \frac{1}{2}\underset{1\times n}{x'}\underset{n\times 1}{x} + \underset{1\times s}{\lambda'}[\underset{s\times n}{G}\underset{n\times 1}{x} - \underset{s\times s}{(F'F)^{-1}}\underset{s\times m}{F'}\underset{m\times 1}{y}] \to \min_{x,\lambda}$$

$$\left.\begin{aligned}\frac{\partial\mathcal{L}(x,\lambda)}{\partial x}(\hat{x},\hat{\lambda}) &= \underset{n\times n}{I}\underset{n\times 1}{\hat{x}} + \underset{n\times s}{G'}\underset{s\times 1}{\hat{\lambda}} = \underset{n\times 1}{0}\\ \frac{\partial\mathcal{L}(x,\lambda)}{\partial \lambda}(\hat{x},\hat{\lambda}) &= \underset{s\times n}{G}\underset{n\times 1}{\hat{x}} - \underset{s\times s}{(F'F)^{-1}}\underset{s\times m}{F'}\underset{m\times 1}{y} = \underset{s\times 1}{0}\end{aligned}\right\} \Rightarrow \begin{bmatrix}\underset{n\times n}{I} & \underset{n\times s}{G'}\\ \underset{s\times n}{G} & \underset{s\times s}{0}\end{bmatrix}\begin{bmatrix}\underset{n\times 1}{\hat{x}}\\ \underset{s\times 1}{\hat{\lambda}}\end{bmatrix} = \begin{bmatrix}\underset{n\times 1}{0}\\ \underset{s\times s}{(F'F)^{-1}}\underset{s\times m}{F'}\underset{m\times 1}{y}\end{bmatrix}$$

or

$$\hat{x} = \underbrace{\underset{n\times s}{G'}(\underset{\substack{s\times n\\ \Downarrow\\ \text{full rank matrix}}}{G}\underset{n\times s}{G'})^{-1}}_{\text{"right inverse of G"}}\underbrace{(\underset{\substack{s\times m\\ \Downarrow\\ \text{full rank matrix}}}{F'}\underset{m\times s}{F})^{-1}\underset{s\times m}{F'}}_{\text{"left inverse of F"}}y = A^{+}y = (A'A)^{+}A'y = N^{+}A'y$$

A^{+} is the "pseudoinverse" of A

→ A-model: Free datum

A-model: Free datum

Solution approach 3): Singular value decomposition of A and pseudoinverse of N

Any rectangular m×n matrix A with rank s can be diagonalized using the singular value decomposition (svd) in such a way that $\underset{m\times n}{A}\,\underset{n\times n}{V} = \underset{m\times m}{U}\,\underset{m\times n}{\Sigma}$, $\underset{m\times n}{A} = \underset{m\times m}{U}\,\underset{m\times n}{\Sigma}\,\underset{n\times n}{V'}$.

U and V are orthogonal matrices, Σ is the m×n matrix of rank s containing the singular values. By computing the normal equation matrix N we get

$$\underset{n\times n}{N} = A'A = V\Sigma'\underbrace{U'U}_{I_m}\Sigma V' = \underset{n\times n}{V}\,\underset{n\times n}{\Lambda}\,\underset{n\times n}{V'} \;,\quad \underset{n\times n}{\Lambda} = \Sigma'\Sigma \;,$$

where Λ is a diagonal matrix of rank s containing the eigenvalues of N. The n−s zero eigenvalues indicate the datum defect d.

$$\underset{n\times n}{\Lambda} = \begin{bmatrix} \lambda_1 > 0 & & & & & \\ & \ddots & & & \mathbf{0} & \\ & & \lambda_{s=n-d} > 0 & & & \\ & & & \lambda_{s+1} = 0 & & \\ & \mathbf{0} & & & \ddots & \\ & & & & & \lambda_n = 0 \end{bmatrix} = \begin{bmatrix} \underset{s\times s}{\Lambda_1} & 0 \\ 0 & \underset{d\times d}{\Lambda_2} = 0 \end{bmatrix}$$

Note: λ's are eigen-values, not Lagrange multipliers !

$$\hat{x} = N^{+}A'y = V\begin{bmatrix} \underset{s\times s}{\Lambda_1^{-1}} & 0 \\ 0 & \underset{d\times d}{0} \end{bmatrix} V'A'y = V_1\Lambda_1^{-1}V_1'A'y \;,\quad \underset{n\times n}{V} = [\underset{n\times s}{V_1}, \underset{n\times d}{V_2}]$$

MATLAB:
N^{+}=pinv(N)

→ A-model: Free datum

A-model: Free datum

Solution approach 4): Nullspace of N and pseudoinverse N^+

In solution approach 3 matrix $\underset{n\times d}{V_2}$ was obtained. It contains the (orthonormalized) eigenvectors corresponding to the zero eigenvalues of N, forms a basis for the nullspace of N and is used for the computation of the pseudoinverse N^+:

$$\text{(a)}\quad N^+ = (N + V_2V_2')^{-1} - V_2(V_2'V_2V_2'V_2)^{-1}V_2' = (N + V_2V_2')^{-1} - V_2V_2'$$

$$= (N + V_2V_2')^{-1}(I - V_2V_2')$$

$$\text{(b)}\quad \begin{bmatrix} N & V_2 \\ V_2' & 0 \end{bmatrix}^{-1} = \begin{bmatrix} N^+ & V_2(V_2'V_2)^{-1} \\ (V_2'V_2)^{-1}V_2' & 0 \end{bmatrix} = \begin{bmatrix} N^+ & V_2 \\ V_2' & 0 \end{bmatrix}$$

Remark 1: Matrices D (from approach 1) and V_2 can be mutually transformed into each other.

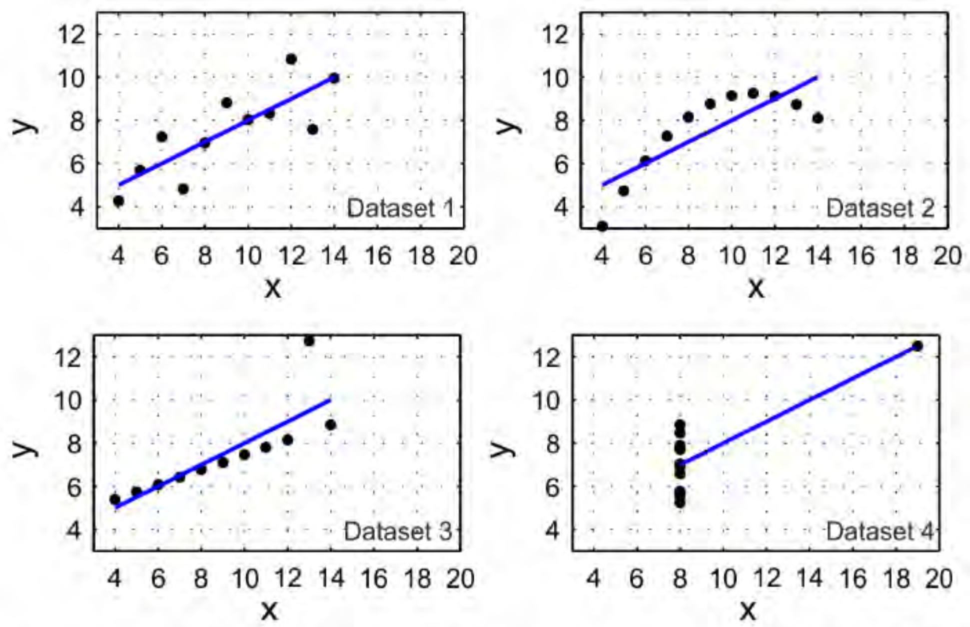
Remark 2: Other approaches, e.g. using the nullspace of A exist.

A numerical example will be given in chapter "Weighted Least Squares: Example free network adjustment".

→ A-model: Linear Regression

A-model: Linear Regression

Regression describes a relationship between one or more input signals and an output signal. If the relation is **linear** in terms of the **parameters** to be estimated from the regression model it is called **linear regression**. The general model for "multiple linear regression" is $y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \ldots + b_u x_{ui} + e_i \;, i = 1, \ldots, m.$ The independent variables x_{1i}, x_{2i}, …, x_{ui} (input signals) can be non-linear functions from other input quantities, y_i is the output signal. e_i are inconsistencies, b_0, b_1, …,b_u are n=u+1 unknown coefficients, to be estimated from m observations y_i.



F.J. Anscombe (1973): Graphs in Statistical Analysis. The American Statistician 27, pp. 17-21

1D-case (u=1, "straight line"):

$$y_i = b_0 + b_1 x_{1i} + e_i \quad \text{or}$$
$$y_i = a + b x_i + e_i$$

Examples:

x: income, y: income tax

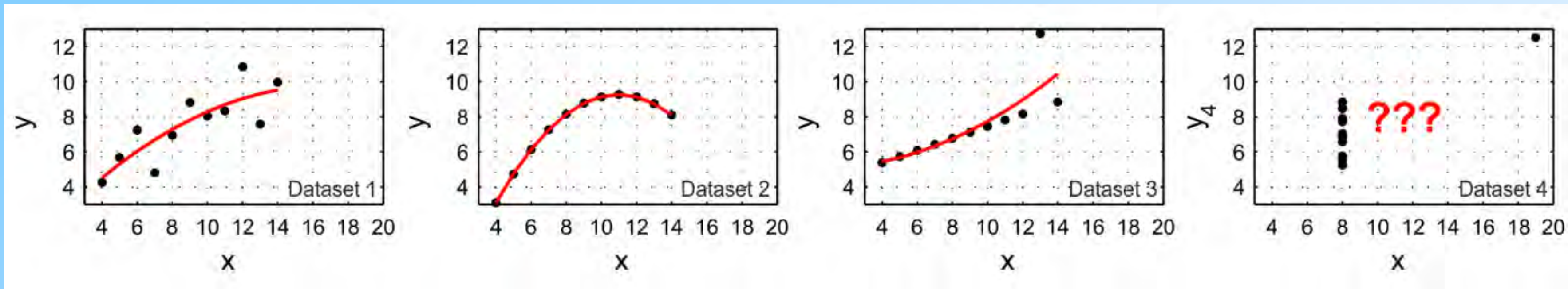
x: time, y: population

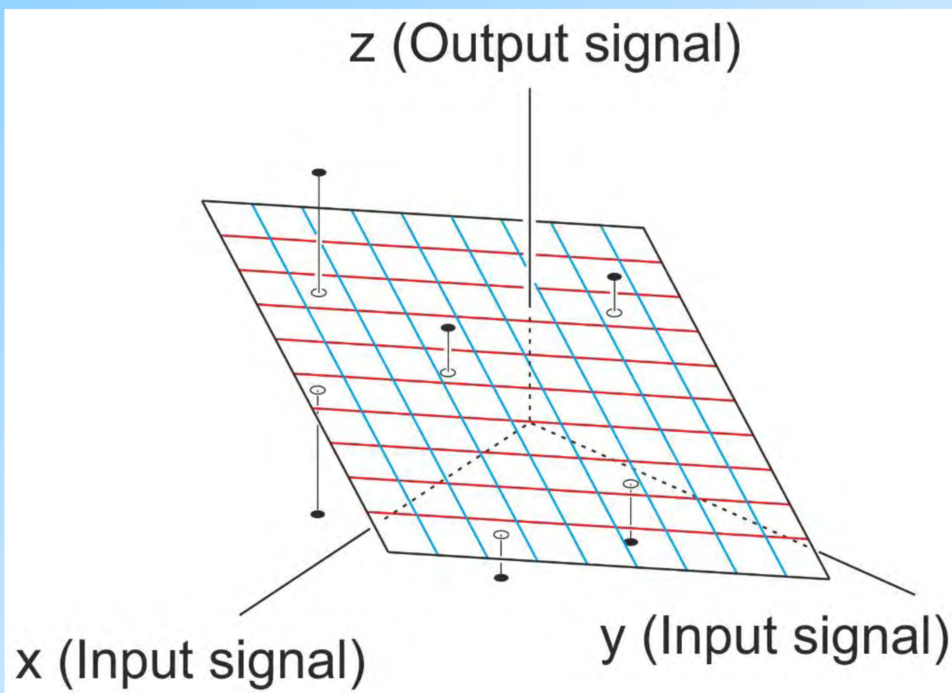
x: resistance, y: voltage (Ohm's law)

→ A-model: Linear Regression

A-model: Linear Regression

1D-case (u=2, "parabola"): $y_i = b_0 + b_1 x_i + b_2 x_i^2 + e_i$





2D-case (u=2, "plane"):

$$y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + e_i$$

or

$$z_i = a + bx_i + cy_i + e_i$$

Example:
x, y: coordinates
y: Heights

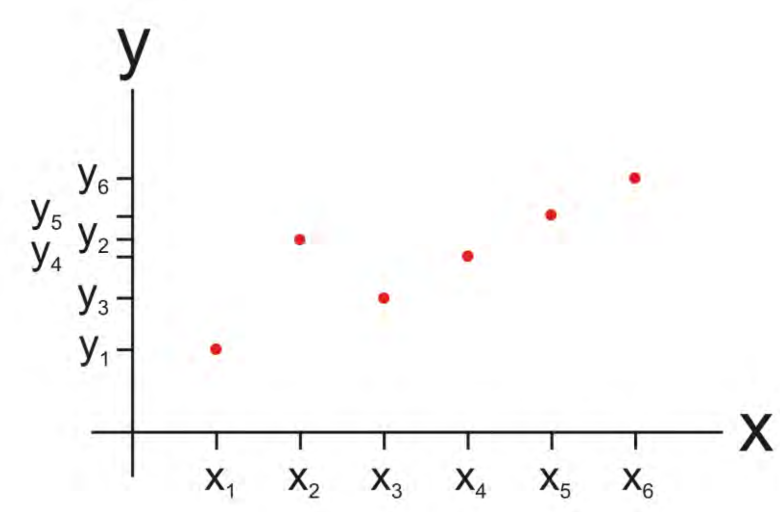
NE

→ A-model: Linear Regression

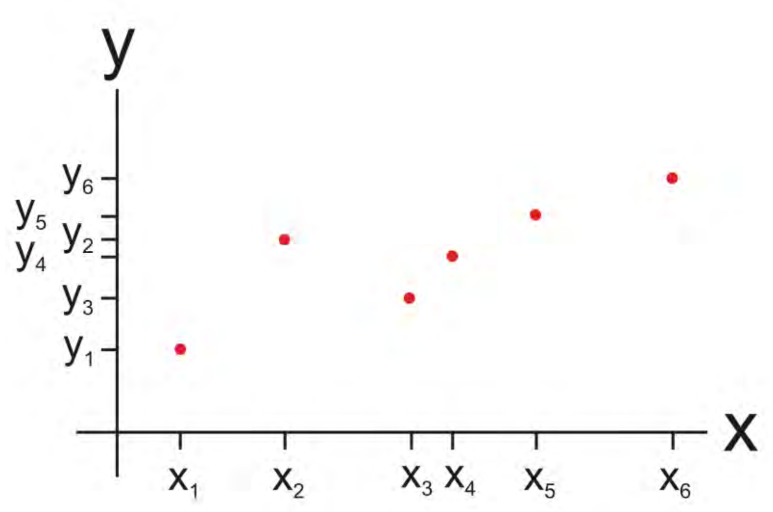
A-model: Linear Regression

The subject of linear regression is to find, e.g., constants $\hat{a}$ (axis intercept) and $\hat{b}$ (slope) as estimates for the model parameter a and b, given supporting points x_i on the abscissa (input signal) and measured observations y_i on the ordinate (output signal).

In other words: How can we fit a line to the set of values x_i and y_i ?



regular x spacing



irregular x spacing

NE

→ A-model: Linear Regression