COMP0130 Robot Vision and Navigation

1B: Introduction to Least-Squares Estimation

Dr Paul D Groves

Session Objectives

Show how to

- Use least-squares estimation to determine unknown parameters from a set of measurements
- Extend least-squares estimation to nonlinear problems
- Account for variation in measurement quality in a least-squares solution

Apply these techniques to some example problems

Contents

1. Formulating the Problem
2. Linear Least-Squares Estimation
3. Applying Least Squares to Nonlinear Problems
4. Weighted Least-Squares Estimation

1. Formulating the Problem

Mathematical Notation

$a,\ A$ Italic for scalars

$\mathbf{a}$ Bold, lower-case for vectors

$\mathbf{A}$ Bold capitals for matrices

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix}$$

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$$

$$\mathbf{a} = \begin{pmatrix} a_1 & a_2 & \cdots & a_m \end{pmatrix}^{\mathrm{T}}$$

Different symbols mean different things

Non-standard notation is occasionally used to avoid clashes

1. Formulating the Problem

The Problem (1)

We want to build a mathematical model from experimental data

Suppose z is a function of y:

$$z = G(y)$$

where G is an unknown function

If we have some pairs of observations:

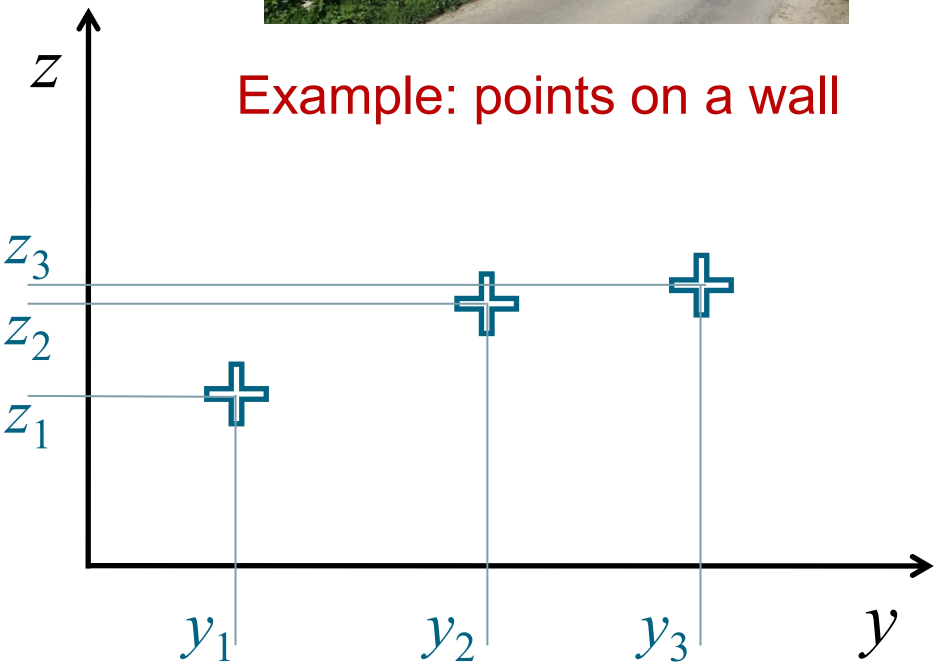
y_1, z_1 $[z_1 = G(y_1)]$

y_2, z_2 $[z_2 = G(y_2)]$

y_3, z_3 $[z_3 = G(y_3)]$

How do we find G ?

Example: points on a wall

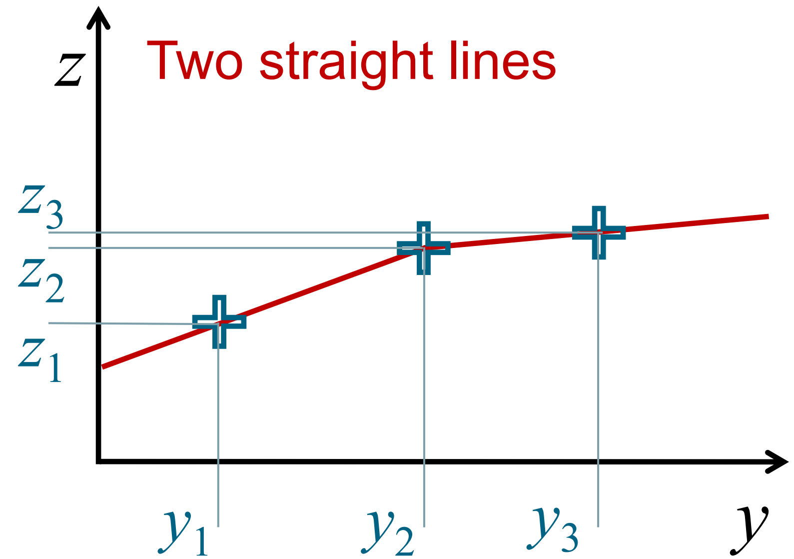


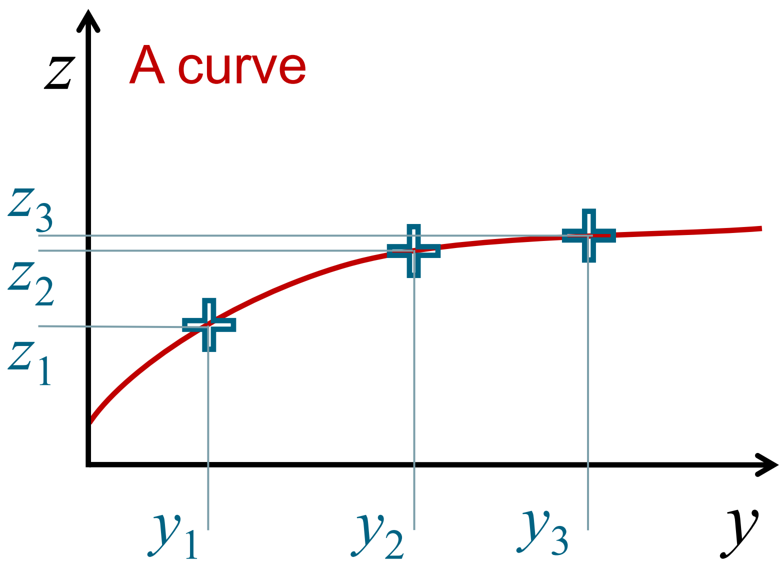
1. Formulating the Problem

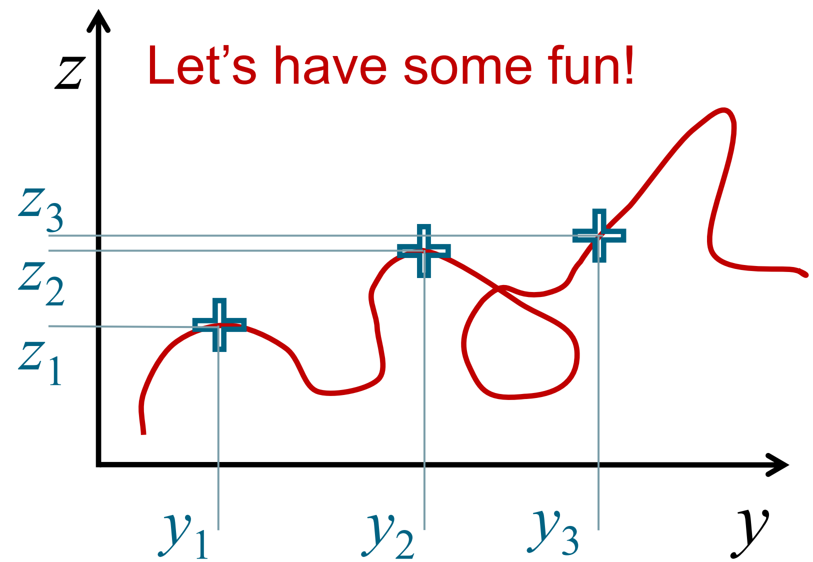
The Problem (2)

$$z = G(y)$$

There's lots of options for G:







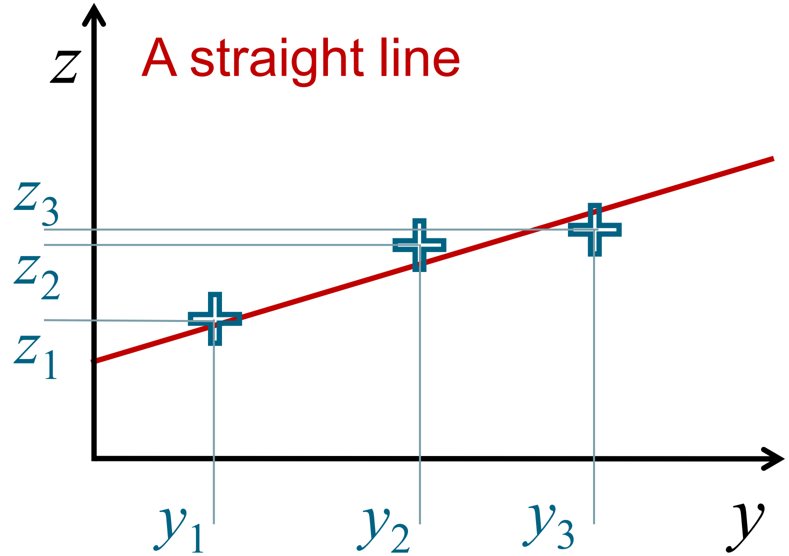
1. Formulating the Problem

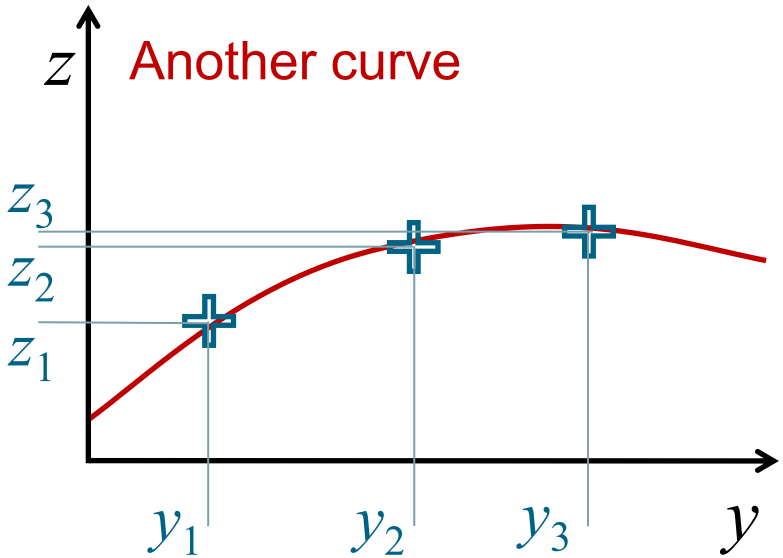
The Problem (3)

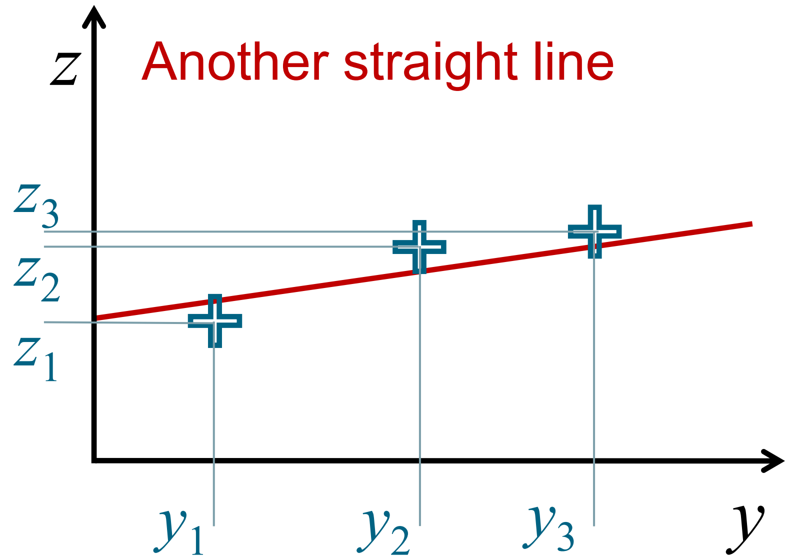
$$z = G(y)$$

There's lots of options for G:

Even more if you assume the observations have errors:

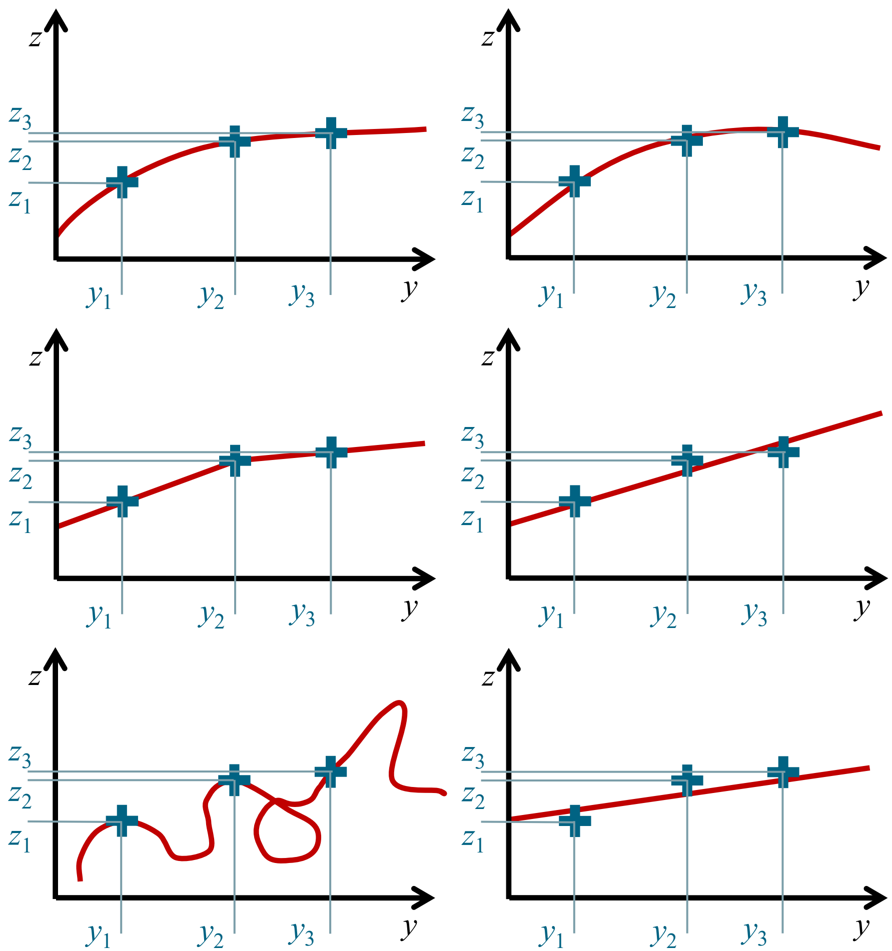






1. Formulating the Problem

The Problem (4)



We need to know what the shape of the function should be

We use the physics of the problem to propose a suitable model

e.g. a wall forms a straight line in the horizontal plane

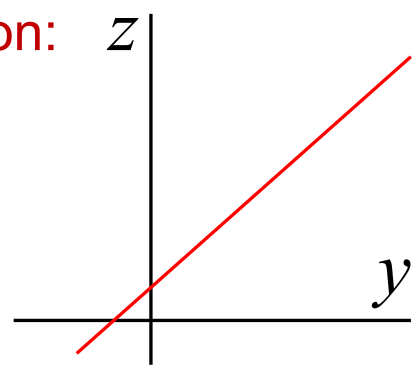
1. Formulating the Problem

Assume the shape of the function

We can select a suitable model based on knowledge of the physics:

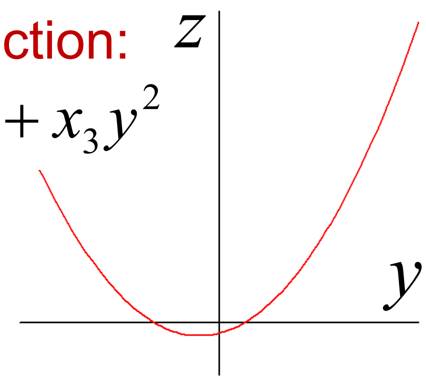
Linear function:

$$z = x_1 + x_2 y$$



Quadratic function:

$$z = x_1 + x_2 y + x_3 y^2$$



Fourier series: $$z = x_1 \cos y + x_2 \sin y + x_3 \cos 2y + x_4 \sin 2y + \ldots$$

Now only the coefficients need to be determined:

$$\mathbf{x} = \begin{pmatrix} x_1 & x_2 & \cdots \end{pmatrix}^{\mathrm{T}}$$

Which greatly simplifies the problem

1. Formulating the Problem

The Measurement Model

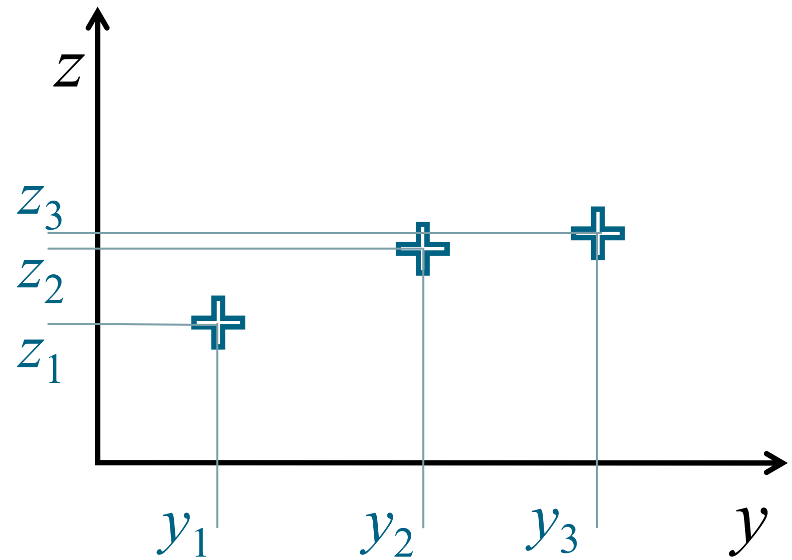
With the shape of the function known,

$$z = G(y) = h(\mathbf{x}, y)$$

where h is a known function and

$$\mathbf{x} = \begin{pmatrix} x_1 & x_2 & \cdots \end{pmatrix}^{\mathrm{T}}$$

This is a *measurement model*



It expresses a known observation or *measurement*, z, in terms of

- another known observation, y,
- the unknown coefficients of the model, $\mathbf{x}$

The coefficients, $\mathbf{x}$, are known as **states** or parameters

The function h is known as the **measurement function**

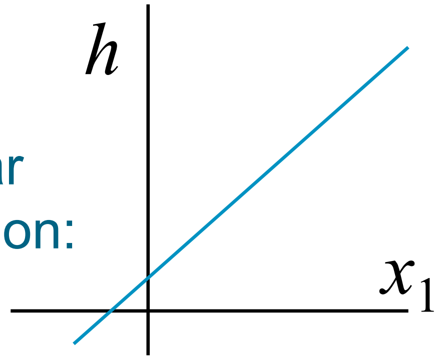
1. Formulating the Problem

Linear Measurement Models

In general,

$$z = h(\mathbf{x}, y)$$

Linear function:



If z is a linear function of the **all** of the coefficients, $\mathbf{x}$, then we may write

$$z = \mathbf{H}(y)\mathbf{x} = H_1(y)x_1 + H_2(y)x_2 + \cdots$$

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \end{pmatrix}$$

where $\mathbf{H}$ is the **measurement** or **observation** or **design** matrix, which

- relates the measurements to the states
- is a known matrix function of y, that need not be linear.
- is not a function of $\mathbf{x}$ when h is a linear function of $\mathbf{x}$

1. Formulating the Problem

The Measurement Matrix

In general,

$$z = h(\mathbf{x}, y)$$

For both linear and nonlinear functions of the coefficients, $\mathbf{x}$, the **measurement matrix**, $\mathbf{H}$, comprises the partial derivatives of the measurement function, h, with respect to the states

$$\mathbf{H}(y) = \frac{dz(\mathbf{x}, y)}{d\mathbf{x}} = \frac{dh(\mathbf{x}, y)}{d\mathbf{x}} = \begin{array}{c} \begin{array}{cccc} x_1 & x_2 & & x_n \end{array} \\ \begin{pmatrix} \dfrac{\partial h}{\partial x_1} & \dfrac{\partial h}{\partial x_2} & \cdots & \dfrac{\partial h}{\partial x_n} \end{pmatrix} \end{array}$$

There is one column of $\mathbf{H}$ for each component of the state vector, $\mathbf{x}$

1. Formulating the Problem

Measurement Model of a Line Function

Suppose z is a linear function of y:

$$z = x_1 + x_2 y$$

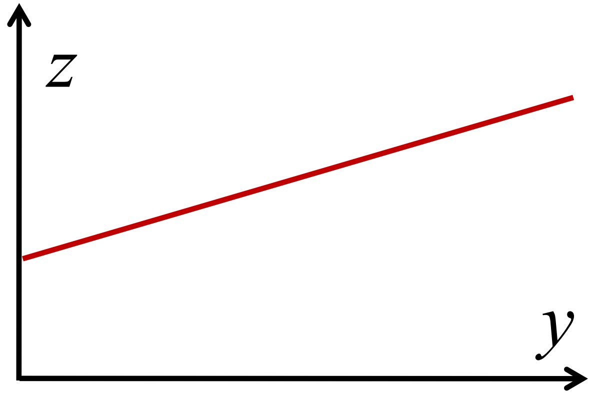
where x_1 is the intercept and x_2 is the gradient

As z is also a linear function of the coefficients, we can write this as:

$$z = \begin{pmatrix} 1 & y \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

or $$z = \mathbf{H}(y)\mathbf{x}$$

where $$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \mathbf{H}(y) = \begin{pmatrix} 1 & y \end{pmatrix}$$



Example: a straight wall

1. Formulating the Problem

Modelling Multiple Measurements

Where the same measurement function, h, applies to multiple measurements:

$$z_1 = h(\mathbf{x}, y_1)$$
$$z_2 = h(\mathbf{x}, y_2)$$
$$\vdots$$

We can write this as:

$$\mathbf{z} = \mathbf{h}(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} h(\mathbf{x}, y_1) \\ h(\mathbf{x}, y_2) \\ \vdots \end{pmatrix}$$

where

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \end{pmatrix}$$

Where $\mathbf{z}$ is a linear function of the **all** of the coefficients, $\mathbf{x}$:

$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \end{pmatrix} = \mathbf{z} = \mathbf{H}(\mathbf{y})\mathbf{x} = \begin{pmatrix} \mathbf{H}'(y_1) \\ \mathbf{H}'(y_2) \\ \vdots \end{pmatrix} \mathbf{x}$$

1. Formulating the Problem

Matrix Solution of Linear Equations

For a set of linear equations written in matrix-vector form as

$$\mathbf{z} = \mathbf{H}\mathbf{x}$$

Where the number of equations equals the number of unknowns, $\mathbf{H}$ is square so generally has an inverse.

We can multiply both sides of the equation by this, giving

$$\mathbf{H}^{-1}\mathbf{z} = \mathbf{H}^{-1}\mathbf{H}\mathbf{x}$$

Multiplying a matrix by its inverse gives the identity matrix, so

$$\mathbf{H}^{-1}\mathbf{H} = \mathbf{I} = \begin{pmatrix} 1 & 0 & \cdots \\ 0 & 1 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Re-arranging, $\mathbf{x} = \mathbf{H}^{-1}\mathbf{z}$

BUT…This only works when $\mathbf{H}$ is **square** and **nonsingular** (i.e., $|\mathbf{H}| \neq 0$)

1. Formulating the Problem

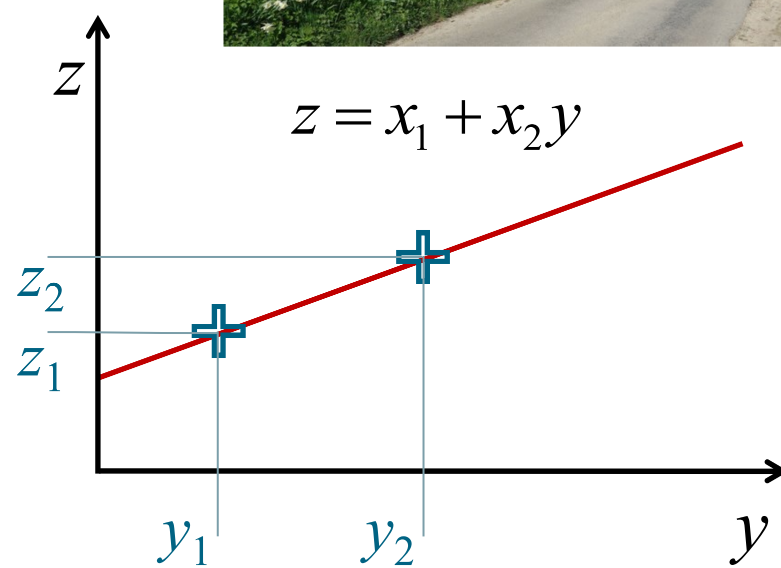
Example 1: A Straight Line Function (1)

At least two y, z observations are needed to solve for x_1 and x_2:

$$\mathbf{z} = \mathbf{h}(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} h(\mathbf{x}, y_1) \\ h(\mathbf{x}, y_2) \end{pmatrix} = \begin{pmatrix} x_1 + x_2 y_1 \\ x_1 + x_2 y_2 \end{pmatrix}$$

where:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad \mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$



As $\mathbf{z}$ is a linear function of both x_1 and x_2

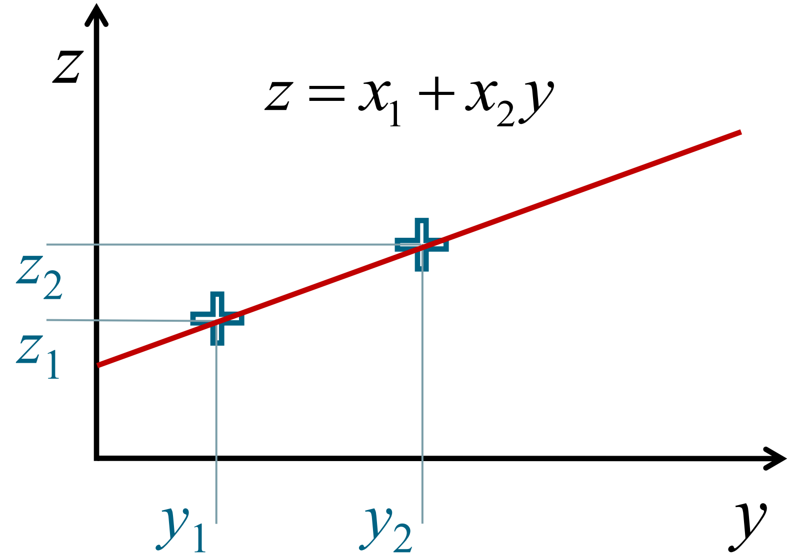
$$\mathbf{z} = \mathbf{H}(\mathbf{y})\mathbf{x} = \begin{pmatrix} 1 & y_1 \\ 1 & y_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

1. Formulating the Problem

Example 1: A Straight Line Function (2)

We are solving $\mathbf{z} = \mathbf{H}\mathbf{x}$ where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \quad \mathbf{H}(\mathbf{y}) = \begin{pmatrix} 1 & y_1 \\ 1 & y_2 \end{pmatrix}$$



The measurement matrix, $\mathbf{H}$, is square and non-singular (provided $y_1 \neq y_2$), so it can be inverted.

The solution is thus $\mathbf{x} = \mathbf{H}^{-1}\mathbf{z}$

Let $(y_1, z_1) = (4, 4)$ and $(y_2, z_2) = (12, 6)$ Therefore:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & y_1 \\ 1 & y_2 \end{pmatrix}^{-1} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 1 & 12 \end{pmatrix}^{-1} \begin{pmatrix} 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 1.5 & -0.5 \\ -0.125 & 0.125 \end{pmatrix} \begin{pmatrix} 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 3 \\ 0.25 \end{pmatrix}$$

See *RVN Least-Squares Examples.xlsx* on Moodle

1. Formulating the Problem

General Problem Formulation

A measurement, z_1, can depend on multiple known parameters, $y_1, y_2 \ldots$

A known parameter, y_1, can impact multiple measurements, $z_1, z_2 \ldots$

Any component of $\mathbf{z}$ and $\mathbf{h}$ can be a function of **any** component of $\mathbf{y}$.

Different components of $\mathbf{z}$ and $\mathbf{h}$ can also be functions of different states

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \end{pmatrix} \quad \mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \end{pmatrix}$$

$$\mathbf{z} = \mathbf{h}(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} h_1(\mathbf{x}, \mathbf{y}) \\ h_2(\mathbf{x}, \mathbf{y}) \\ \vdots \end{pmatrix}$$

$$\mathbf{H}(\mathbf{y}) = \begin{pmatrix} \dfrac{\partial h_1(\mathbf{x}, \mathbf{y})}{\partial x_1} & \dfrac{\partial h_1(\mathbf{x}, \mathbf{y})}{\partial x_2} & \cdots \\ \dfrac{\partial h_2(\mathbf{x}, \mathbf{y})}{\partial x_1} & \dfrac{\partial h_2(\mathbf{x}, \mathbf{y})}{\partial x_2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

jacobian Matrix

描述函数 Oupt 对于 Inpt 的变化

1. Formulating the Problem

Handling Real Measurements (1)

Measurements are always subject to error

$$\underbrace{\tilde{z}}_{\text{Measured value}} = \underbrace{z}_{\text{True value}} + \underbrace{\varepsilon}_{\text{Error}}$$

~ is called ‘tilde’

Therefore, states or parameters determined from those measurements are also subject to error

$$\underbrace{\hat{x}}_{\text{Estimated value}} = \underbrace{x}_{\text{True value}} + \underbrace{e}_{\text{Error}}$$

^ is called ‘caret’

For a linear system, if $\hat{\mathbf{x}} = \mathbf{H}^{-1}\tilde{\mathbf{z}}$ and $\mathbf{x} = \mathbf{H}^{-1}\mathbf{z}$ then $\mathbf{e} = \mathbf{H}^{-1}\boldsymbol{\varepsilon}$

1. Formulating the Problem

Handling Real Measurements (2)

Because measurements are always subject to error, states estimated from those measurements will also be subject to error

The effect of ***random*** errors can be reduced by using more measurements

Solution error standard deviation

$$\sigma_e \propto \frac{1}{\sqrt{m}}$$

Is proportional to

Number of measurements

But, we cannot use $\mathbf{x} = \mathbf{H}^{-1}\mathbf{z}$ if there are more measurements than states

1. Only square matrices can be inverted
2. The simultaneous equations will contradict each other because of the measurement errors

We need a new approach: **Least-squares Estimation**

Contents

1. Formulating the Problem
2. Linear Least-Squares Estimation
3. Applying Least Squares to Nonlinear Problems
4. Weighted Least-Squares Estimation

2. Linear Least-Squares Estimation

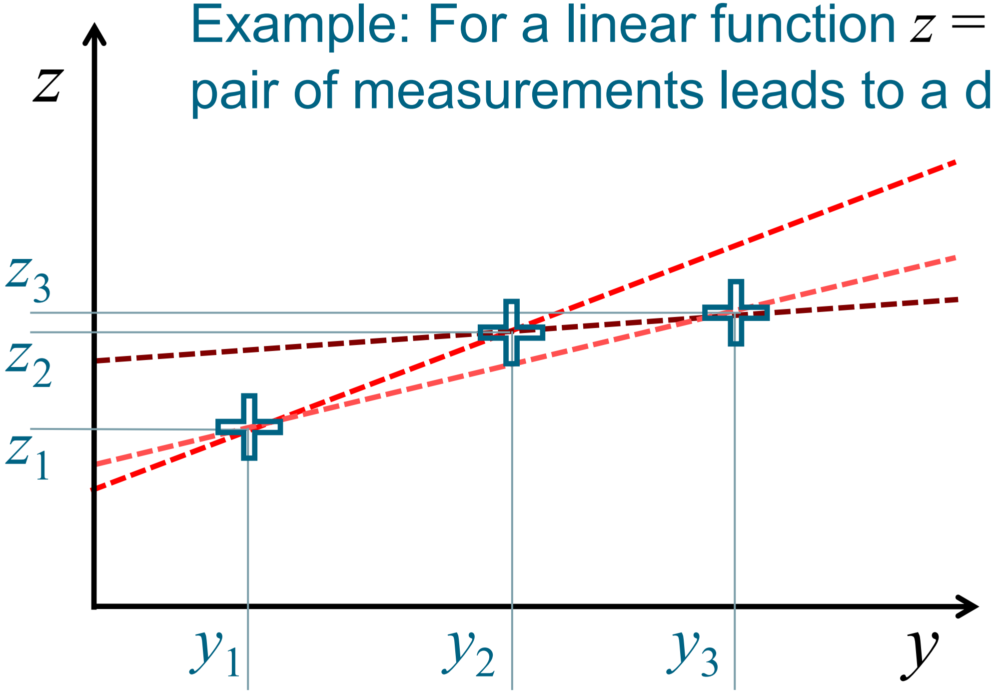
More Measurements than States

Due to measurement errors, observations will contradict each other

Different combinations of measurements give different solutions

There is no exact solution

Example: For a linear function $z = x_1 + x_2 y$ – each pair of measurements leads to a different solution



Formulating the problem as $\mathbf{z} = \mathbf{H}(\mathbf{y})\mathbf{x}$

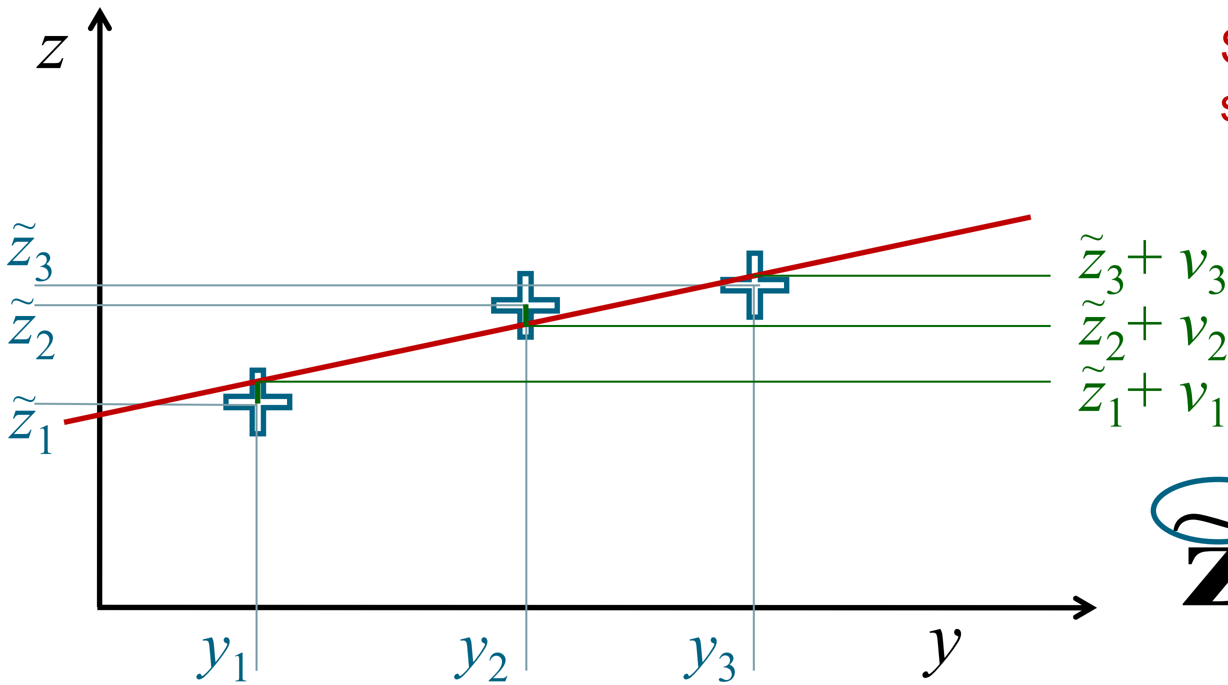
$\mathbf{H}$ cannot be inverted

What do we do?

Adjusting the Measurements to Fit

We assume that z is subject to measurement error, but ignore errors in y

We make an adjustment to each z observation to make $\mathbf{z}$ fit the function $\mathbf{h}(\mathbf{x},\mathbf{y})$. This adjustment is called the residual, v.



Sometimes the opposite sign convention is used:

$v = -\,\delta z^{+}$

$\tilde{\mathbf{z}}$

Tilde denotes a measurement – not the exact value

Modifying the Measurement Model

General measurement model: $\tilde{\mathbf{z}} + \mathbf{v} = \mathbf{h}(\mathbf{x}, \mathbf{y})$

Linear measurement model: $\tilde{\mathbf{z}} + \mathbf{v} = \mathbf{H}(\mathbf{y})\mathbf{x}$

where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \quad \tilde{\mathbf{z}} = \begin{pmatrix} \tilde{z}_1 \\ \tilde{z}_2 \\ \vdots \\ \tilde{z}_m \end{pmatrix} \quad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix}$$

n = number of states

m = number of measurements

$$\mathbf{H}(\mathbf{y}) = \begin{pmatrix} \partial h_1/\partial x_1 & \partial h_1/\partial x_2 & \cdots & \partial h_1/\partial x_n \\ \partial h_2/\partial x_1 & \partial h_2/\partial x_2 & \cdots & \partial h_2/\partial x_n \\ \vdots & \vdots & \ddots & \vdots \\ \partial h_m/\partial x_1 & \partial h_m/\partial x_2 & \cdots & \partial h_m/\partial x_n \end{pmatrix}$$

How do we solve this?

2. Linear Least-Squares Estimation

Obtaining a Solution

Known

$$\boxed{\tilde{\mathbf{z}}} + \boxed{\mathbf{v}} = \boxed{\mathbf{H}(\mathbf{y})}\boxed{\mathbf{x}}$$

Unknown

$$\begin{pmatrix} \tilde{z}_1 \\ \tilde{z}_2 \\ \vdots \\ \tilde{z}_m \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} & \cdots & H_{1n} \\ H_{21} & H_{22} & \cdots & H_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ H_{m1} & H_{m2} & \cdots & H_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

There are as many residuals as rows in the equation (m)

$\therefore$ There are more unknown terms $(m+n)$ than simultaneous equations (m)

The problem is underdetermined

$\therefore$ There is no unique solution for states, $\mathbf{x}$, and residuals, $\mathbf{v}$

$\therefore$ We need more information

Introducing the Least-Squares Constraint

Known

$$\boxed{\tilde{\mathbf{z}}} + \boxed{\mathbf{v}} = \boxed{\mathbf{H}(\mathbf{y})}\boxed{\mathbf{x}}$$

Unknown

$$\begin{pmatrix} \tilde{z}_1 \\ \tilde{z}_2 \\ \vdots \\ \tilde{z}_m \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} & \cdots & H_{1n} \\ H_{21} & H_{22} & \cdots & H_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ H_{m1} & H_{m2} & \cdots & H_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

We need more information to solve this

The Least-Squares solution is that which minimises the sum of the squares of the residuals

$$\sum_i v_i^2 = \mathbf{v}^{\mathrm{T}}\mathbf{v}$$

It delivers the solution that passes closest to the set of $\mathbf{y}$, $\mathbf{z}$ observations

2. Linear Least-Squares Estimation

Deriving the Linear Least-Squares Solution (1)

To solve for $\mathbf{x}$ and $\mathbf{v}$

$$\tilde{\mathbf{z}} + \mathbf{v} = \mathbf{H}(\mathbf{y})\hat{\mathbf{x}}^{+} \qquad - \ (1)$$

Constraint: select values of $\mathbf{x}$ that minimise the sum of squares of the residuals, $\sum_i v_i^2$

Thus…

$$\frac{\partial}{\partial \hat{\mathbf{x}}^{+}}\left(\mathbf{v}^{\mathrm{T}}\mathbf{v}\right) = \mathbf{0} \qquad - \ (2)$$

Carat denotes an estimated value – solution is not exact

$$\hat{\mathbf{x}}^{+}$$

"+" denotes 'a posteriori' – incorporating the measurement data

Substituting (1) into (2) :

$$\frac{\partial}{\partial \hat{\mathbf{x}}^{+}}\left[\left(\mathbf{H}(\mathbf{y})\hat{\mathbf{x}}^{+} - \tilde{\mathbf{z}}\right)^{\mathrm{T}}\left(\mathbf{H}(\mathbf{y})\hat{\mathbf{x}}^{+} - \tilde{\mathbf{z}}\right)\right] = \mathbf{0} \qquad - \ (3)$$

2. Linear Least-Squares Estimation

Deriving the Linear Least-Squares Solution (2)

From before: $$\frac{\partial}{\partial \hat{\mathbf{x}}^{+}}\left[\left(\mathbf{H}(\mathbf{y})\hat{\mathbf{x}}^{+}-\tilde{\mathbf{z}}\right)^{\mathrm{T}}\left(\mathbf{H}(\mathbf{y})\hat{\mathbf{x}}^{+}-\tilde{\mathbf{z}}\right)\right]=\mathbf{0} \quad - (3)$$

Expanding: $$\frac{\partial}{\partial \hat{\mathbf{x}}^{+}}\left[\hat{\mathbf{x}}^{+\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{H}\hat{\mathbf{x}}^{+}-\hat{\mathbf{x}}^{+\mathrm{T}}\mathbf{H}^{\mathrm{T}}\tilde{\mathbf{z}}-\tilde{\mathbf{z}}^{\mathrm{T}}\mathbf{H}\hat{\mathbf{x}}^{+}+\tilde{\mathbf{z}}^{\mathrm{T}}\tilde{\mathbf{z}}\right]=\mathbf{0} \quad - (4)$$

Differentiating:

$$2\hat{\mathbf{x}}^{+\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{H}-2\tilde{\mathbf{z}}^{\mathrm{T}}\mathbf{H}=\mathbf{0} \quad - (5)$$

Noting that $\frac{\partial}{\partial \mathbf{a}}\mathbf{a}^{\mathrm{T}}\mathbf{b}=\mathbf{b}^{\mathrm{T}}$

Transposing and rearranging: $$\mathbf{H}^{\mathrm{T}}\mathbf{H}\hat{\mathbf{x}}^{+}=\mathbf{H}^{\mathrm{T}}\tilde{\mathbf{z}} \quad - (6)$$

Deriving the Linear Least-Squares Solution (3)

From before: $$\mathbf{H}^{\mathrm{T}}\mathbf{H}\hat{\mathbf{x}}^{+} = \mathbf{H}^{\mathrm{T}}\tilde{\mathbf{z}} \qquad - (6)$$

Multiplying both sides by $(\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}$:

$$\cancel{\left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{H}}\hat{\mathbf{x}}^{+} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\tilde{\mathbf{z}} \qquad - (7)$$

Cancelling: $$\boxed{\hat{\mathbf{x}}^{+} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\tilde{\mathbf{z}}} \qquad - (8)$$

This is the unweighted least-squares solution for a linear problem

Note that $(\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}$ is the *left pseudo-inverse* of $\mathbf{H}$

See also Derivation 1 in RVN Least-Squares Derivations.docx on Moodle

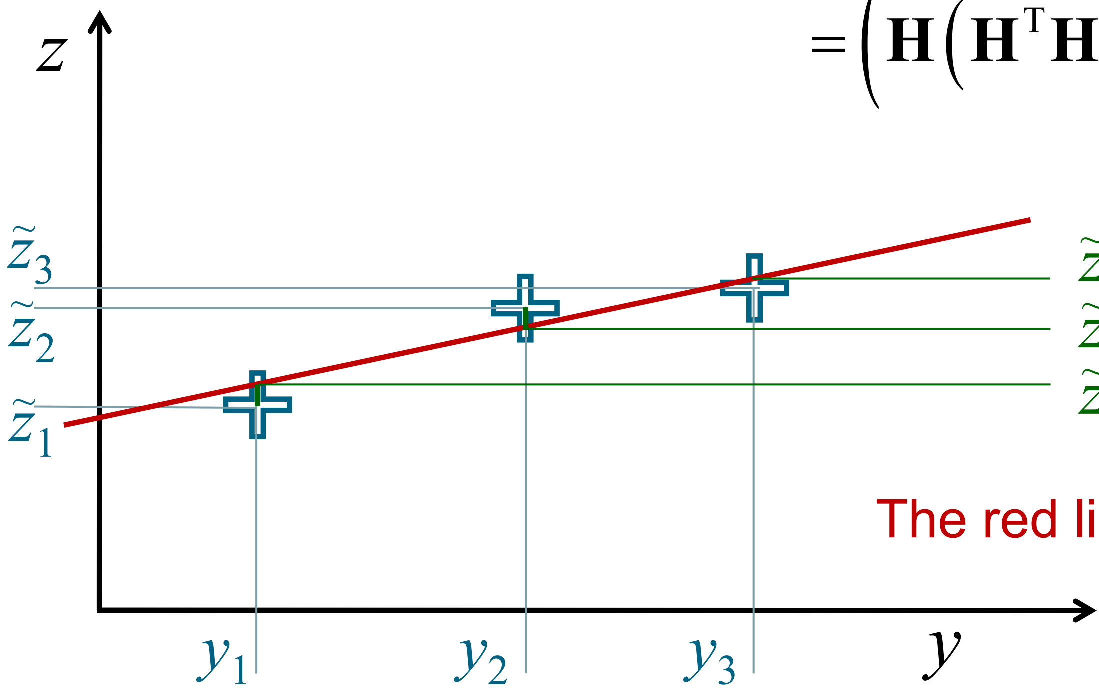
2. Linear Least-Squares Estimation

Residuals

Least-squares solution of $\tilde{\mathbf{z}} + \mathbf{v} = \mathbf{H}(\mathbf{y})\hat{\mathbf{x}}^+$ is $\hat{\mathbf{x}}^+ = \left(\mathbf{H}^\mathrm{T}\mathbf{H}\right)^{-1}\mathbf{H}^\mathrm{T}\tilde{\mathbf{z}}$

The residuals are given by

$$\begin{aligned}\mathbf{v} &= \mathbf{H}\hat{\mathbf{x}}^+ - \tilde{\mathbf{z}} \\ &= \left(\mathbf{H}\left(\mathbf{H}^\mathrm{T}\mathbf{H}\right)^{-1}\mathbf{H}^\mathrm{T} - \mathbf{I}\right)\tilde{\mathbf{z}}\end{aligned}$$



The red line shows $\hat{\mathbf{z}} = \mathbf{H}(\mathbf{y})\hat{\mathbf{x}}^+$

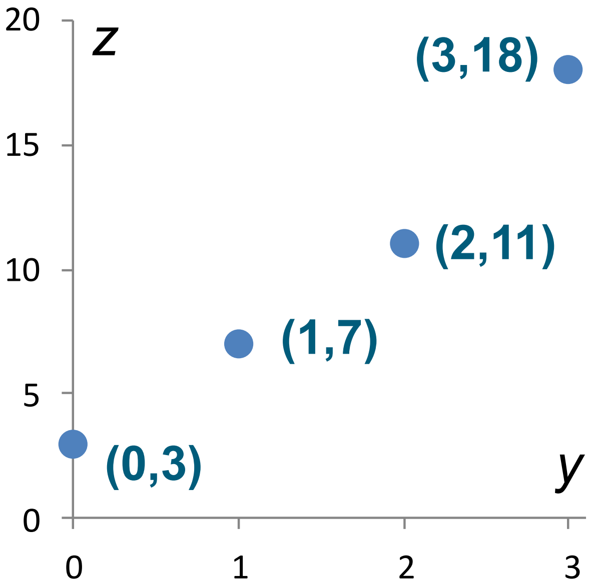
2. Linear Least-Squares Estimation

Example 2: A Straight Line (1)

We have y, z coordinates of four points along a wall. We assume…

1. The y coordinates are exact
2. The z coordinates have measurement errors
3. The Wall is straight

A straight line is represented by $z = x_1 + x_2 y$, where x_1 is the intercept and x_2 is the gradient.



We use least-squares estimation to obtain values of x_1 and x_2 from the data

See *RVN Least-Squares Examples.xlsx* on Moodle

2. Linear Least-Squares Estimation

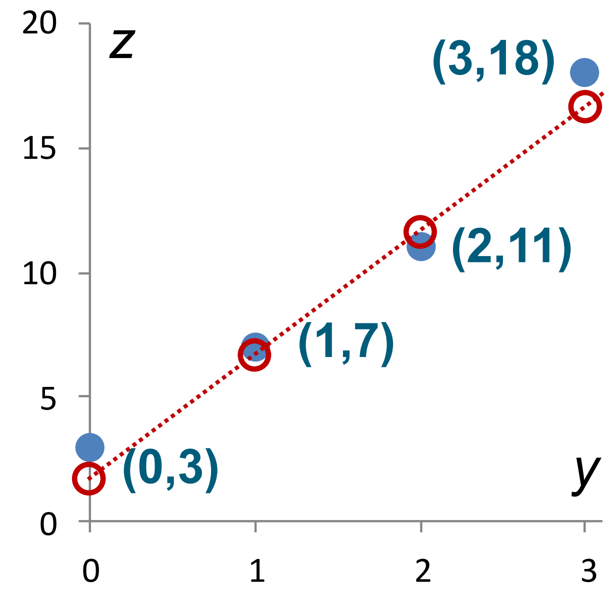
Example 2: A Straight Line (2)

Our model for a straight line is $z = x_1 + x_2 y$

This is linear, so $\mathbf{z} = \mathbf{H}(\mathbf{y})\mathbf{x}$

and $\tilde{\mathbf{z}} + \mathbf{v} = \mathbf{H}(\mathbf{y})\hat{\mathbf{x}}^{+}$ where

$$\tilde{\mathbf{z}} = \begin{pmatrix} \tilde{z}_1 \\ \tilde{z}_2 \\ \tilde{z}_3 \\ \tilde{z}_4 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \\ 11 \\ 18 \end{pmatrix} \quad \mathbf{H}(\mathbf{y}) = \begin{pmatrix} 1 & y_1 \\ 1 & y_2 \\ 1 & y_3 \\ 1 & y_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}$$



The solution is $$\begin{pmatrix} \hat{x}_1^{+} \\ \hat{x}_2^{+} \end{pmatrix} = \hat{\mathbf{x}}^{+} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\tilde{\mathbf{z}} = \begin{pmatrix} 2.4 \\ 4.9 \end{pmatrix} \Rightarrow z = 2.4 + 4.9y$$

See *RVN Least-Squares Examples.xlsx* on Moodle

Contents

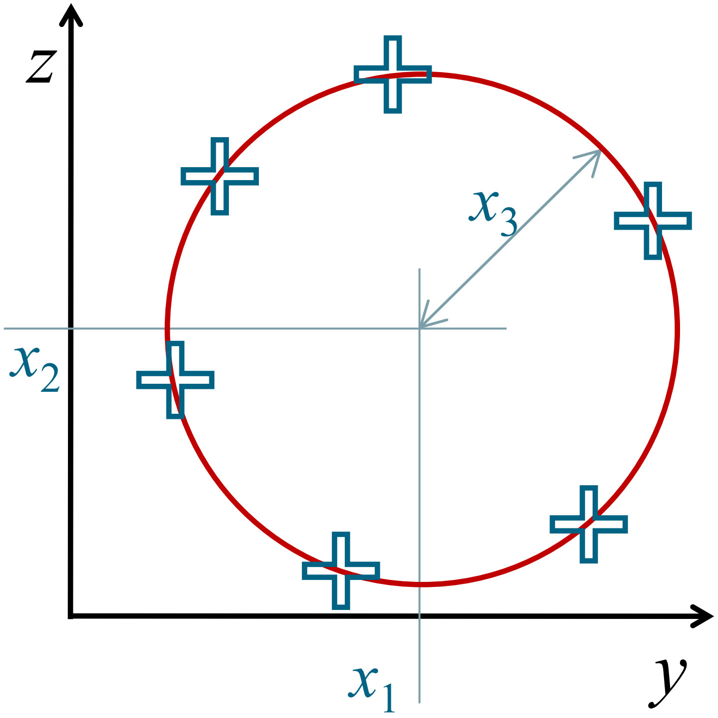
1. Formulating the Problem
2. Linear Least-Squares Estimation
3. Applying Least Squares to Nonlinear Problems
4. Weighted Least-Squares Estimation

3. Applying Least Squares to Nonlinear Problems

Nonlinear Problems (1)

Unfortunately, observations are not always linear functions of the states:

Example A: Finding the centre and radius of a chimney



Applying Pythagoras' theorem:

$$x_3^2 = (y - x_1)^2 + (z - x_2)^2$$

$$\Rightarrow z = x_2 \pm \sqrt{x_3^2 - (y - x_1)^2}$$

z is a linear function of x_2,

But it is a *nonlinear* function of x_1 and x_3

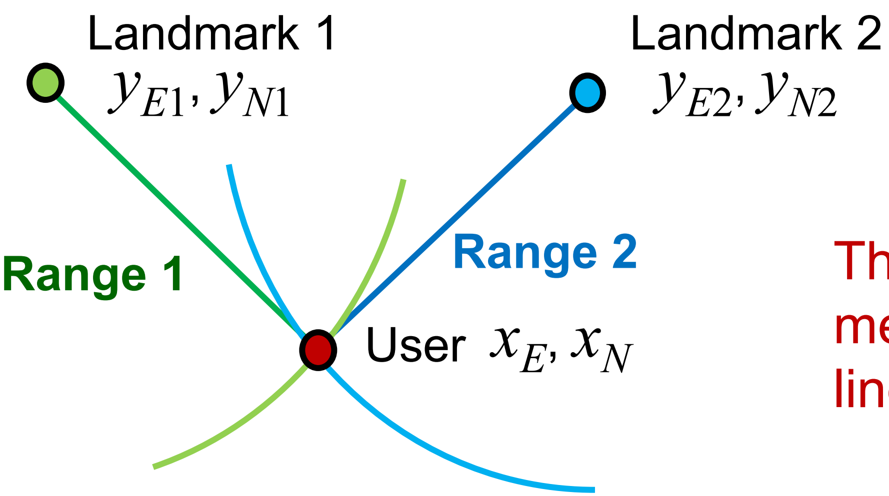
The least-squares method can only solve linear problems

3. Applying Least Squares to Nonlinear Problems

Nonlinear Problems (2)

Unfortunately, observations are not always a linear functions of the states.

Example B: Determining positions from ranging measurements, such as GNSS



The least-squares method can only solve linear problems

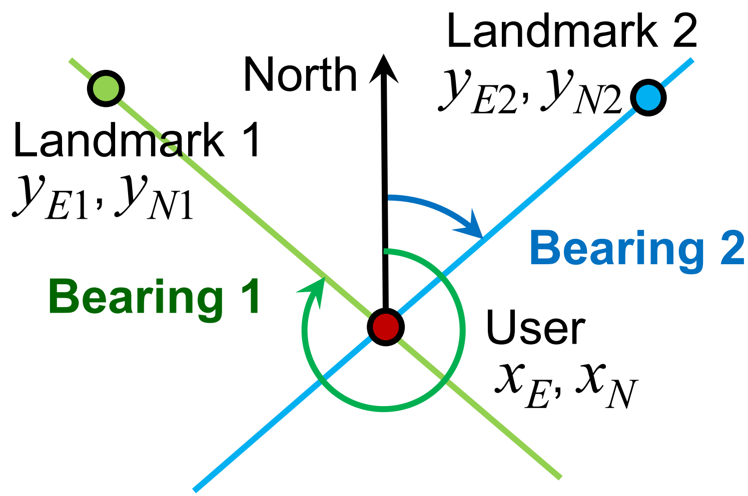
$$\begin{pmatrix} z_{r1} \\ z_{r2} \end{pmatrix} = \begin{pmatrix} h_1(\mathbf{x}, \mathbf{y}) \\ h_2(\mathbf{x}, \mathbf{y}) \end{pmatrix} = \begin{pmatrix} \sqrt{(y_{E1} - x_E)^2 + (y_{N1} - x_N)^2} \\ \sqrt{(y_{E2} - x_E)^2 + (y_{N2} - x_N)^2} \end{pmatrix}$$

3. Applying Least Squares to Nonlinear Problems

Nonlinear Problems (3)

Unfortunately, observations are not always a linear functions of the states.

Example C: Determining positions from optical angle measurements



$$\begin{pmatrix} z_{\psi 1} \\ z_{\psi 2} \end{pmatrix} = \begin{pmatrix} h_1(\mathbf{x}, \mathbf{y}) \\ h_2(\mathbf{x}, \mathbf{y}) \end{pmatrix} = \begin{pmatrix} \arctan_2\left(\left(y_{E1} - x_E\right), \left(y_{N1} - x_N\right)\right) \\ \arctan_2\left(\left(y_{E2} - x_E\right), \left(y_{N2} - x_N\right)\right) \end{pmatrix}$$

The least-squares method can only solve linear problems

3. Applying Least Squares to Nonlinear Problems

Finding an Equivalent Linear Problem

We cannot solve a nonlinear problem using least-squares estimation directly

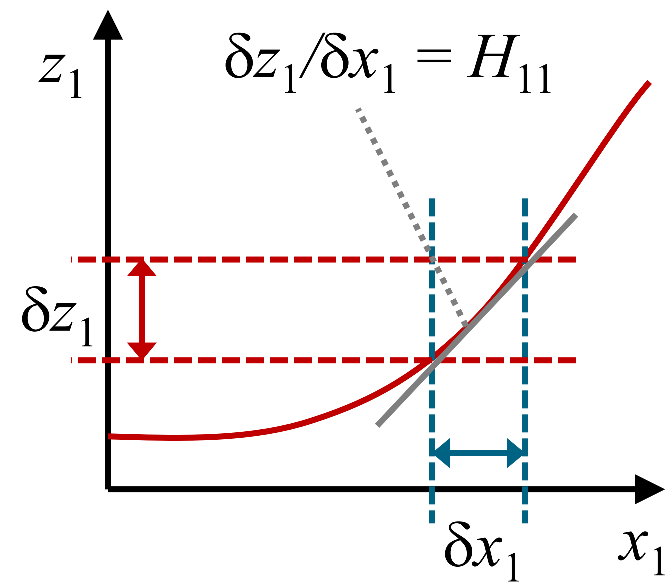
$$\mathbf{z} = \begin{pmatrix} h_1(\mathbf{x},\mathbf{y}) \\ h_2(\mathbf{x},\mathbf{y}) \\ \vdots \\ h_m(\mathbf{x},\mathbf{y}) \end{pmatrix} \equiv \mathbf{h}(\mathbf{x},\mathbf{y}) \neq \mathbf{H}(\mathbf{y})\mathbf{x}$$

Instead, we must formulate an equivalent linear problem, such as

$\delta\mathbf{z} = \mathbf{H}(\mathbf{x},\mathbf{y})\,\delta\mathbf{x}$, where

$\delta\mathbf{z}$ is the change in $\mathbf{z}$, and

$\delta\mathbf{x}$ is the change in $\mathbf{x}$



To use least-squares we must essentially turn a nonlinear problem into a linear one

Linearisation using Taylor's Theorem

Applying **Taylor's theorem** to the measurement model…

$$\mathbf{z} = \mathbf{h}(\mathbf{x},\mathbf{y}) = \mathbf{h}(\mathbf{x}',\mathbf{y}) + \frac{\partial \mathbf{h}(\mathbf{x}',\mathbf{y})}{\partial \mathbf{x}}[\mathbf{x}-\mathbf{x}'] + \boxed{\sum_{r=2}^{\infty} \frac{\partial^r \mathbf{h}(\mathbf{x}',\mathbf{y})}{\partial \mathbf{x}^r} \frac{[\mathbf{x}-\mathbf{x}']^r}{r!}}$$

If we select $\mathbf{x}'$ such that this term is negligible,

then… $$\mathbf{h}(\mathbf{x},\mathbf{y}) \approx \mathbf{h}(\mathbf{x}',\mathbf{y}) + \frac{\partial \mathbf{h}(\mathbf{x}',\mathbf{y})}{\partial \mathbf{x}}[\mathbf{x}-\mathbf{x}']$$

or $\mathbf{h}(\mathbf{x},\mathbf{y}) \approx \mathbf{h}(\mathbf{x}',\mathbf{y}) + \mathbf{H}(\mathbf{x}',\mathbf{y})[\mathbf{x}-\mathbf{x}']$ where $\mathbf{H}(\mathbf{x}',\mathbf{y}) = \dfrac{\partial \mathbf{h}(\mathbf{x}',\mathbf{y})}{\partial \mathbf{x}}$

Rearranging: $$\mathbf{h}(\mathbf{x},\mathbf{y}) - \mathbf{h}(\mathbf{x}',\mathbf{y}) \approx \mathbf{H}(\mathbf{x}',\mathbf{y})[\mathbf{x}-\mathbf{x}']$$

*This first-order approximation is known as **linearisation***

3. Applying Least Squares to Nonlinear Problems

The Measurement Matrix

$\mathbf{H}$ is the **measurement** (or observation matrix), which

- relates **changes** in the measurements to **changes** in the states
- comprises the partial derivatives of $\mathbf{h}$ with respect to the states
- **is a function of both $\mathbf{x}$ and $\mathbf{y}$**

$$\begin{array}{cc} \begin{array}{cccc} x_1 & x_2 & & x_n \end{array} & \\ \mathbf{H}(\mathbf{x}', \mathbf{y}) = \dfrac{\partial \mathbf{h}(\mathbf{x}', \mathbf{y})}{\partial \mathbf{x}} = \left. \begin{pmatrix} \partial h_1/\partial x_1 & \partial h_1/\partial x_2 & \cdots & \partial h_1/\partial x_n \\ \partial h_2/\partial x_1 & \partial h_2/\partial x_2 & \cdots & \partial h_2/\partial x_n \\ \vdots & \vdots & \ddots & \vdots \\ \partial h_m/\partial x_1 & \partial h_m/\partial x_2 & \cdots & \partial h_m/\partial x_n \end{pmatrix} \right|_{\mathbf{x}=\mathbf{x}'} & \begin{array}{c} z_1 \\ z_2 \\ \\ z_m \end{array} \end{array}$$

Each row corresponds to one component of the function, $\mathbf{h}$, *and the measurement,* $\mathbf{z}$

Each column corresponds to one component of the state vector, $\mathbf{x}$

We calculate $\mathbf{H}$ using the predicted values of $\mathbf{x}$. i.e., $\mathbf{x}' = \hat{\mathbf{x}}^-$

3. Applying Least Squares to Nonlinear Problems

Linearising the Problem (1)

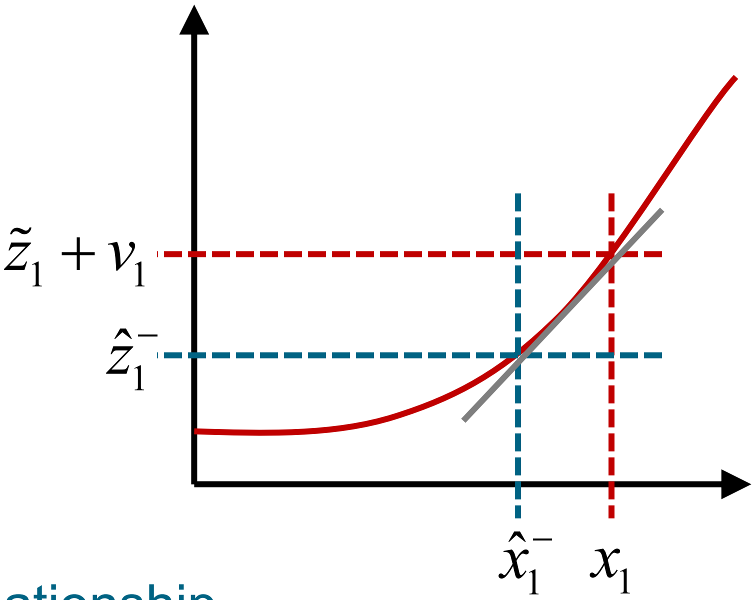
To use least-squares we must turn a nonlinear problem into a linear one

To solve for $\hat{\mathbf{x}}^+$ and $\mathbf{v}$:

$$\tilde{\mathbf{z}} + \mathbf{v} = \mathbf{h}\left(\hat{\mathbf{x}}^+, \mathbf{y}\right) \neq \mathbf{H}\left(\mathbf{y}\right)\hat{\mathbf{x}}^+$$

We use a prediction of the states, $\hat{\mathbf{x}}^-$ to predict the measurements:

$$\hat{\mathbf{z}}^- = \mathbf{h}\left(\hat{\mathbf{x}}^-, \mathbf{y}\right)$$



Subtracting this from both sides:

$$\tilde{\mathbf{z}} - \mathbf{h}\left(\hat{\mathbf{x}}^-, \mathbf{y}\right) + \mathbf{v} = \mathbf{h}\left(\hat{\mathbf{x}}^+, \mathbf{y}\right) - \mathbf{h}\left(\hat{\mathbf{x}}^-, \mathbf{y}\right)$$

We can then use this to model a linear relationship…

Linearising the Problem (2)

From the previous slide…

$$\tilde{\mathbf{z}} - \mathbf{h}\left(\hat{\mathbf{x}}^-, \mathbf{y}\right) + \mathbf{v} = \mathbf{h}\left(\hat{\mathbf{x}}^+, \mathbf{y}\right) - \mathbf{h}\left(\hat{\mathbf{x}}^-, \mathbf{y}\right)$$

$$\approx \mathbf{H}(\hat{\mathbf{x}}^-, \mathbf{y})\left[\hat{\mathbf{x}}^+ - \hat{\mathbf{x}}^-\right]$$

$$= \mathbf{H}(\hat{\mathbf{x}}^-, \mathbf{y})\delta\mathbf{x}$$

where $\delta\mathbf{x} = \hat{\mathbf{x}}^+ - \hat{\mathbf{x}}^-$

$$\mathbf{b} + \mathbf{v} \approx \mathbf{H}(\hat{\mathbf{x}}^-, \mathbf{y})\delta\mathbf{x}$$

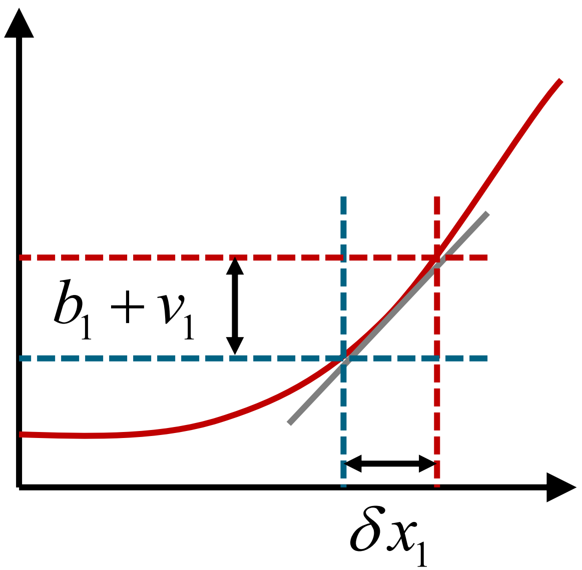
Measurement innovation,

$\mathbf{b}$, or $\delta\mathbf{z}^-$

= measurements minus predictions

First-order Taylor series approximation

– Linearisation



This can be solved using least-squares estimation

3. Applying Least Squares to Nonlinear Problems

Nonlinear Least-Squares Solution

We now have a linear equation to solve for $\delta\mathbf{x}$ and $\mathbf{v}$:

$$\mathbf{b} + \mathbf{v} \approx \mathbf{H}(\hat{\mathbf{x}}^{-}, \mathbf{y})\delta\mathbf{x}$$

We select values of $\delta\mathbf{x}$ that minimise the sum of squares of the residuals, $\sum_i v_i^2$

The solution is the same as for linear least-squares estimation.

Thus: $\delta\mathbf{x} \approx \left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{b}$

Giving
$$\begin{aligned}\hat{\mathbf{x}}^{+} &= \hat{\mathbf{x}}^{-} + \delta\mathbf{x} \\ &\approx \hat{\mathbf{x}}^{-} + \left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{b}\end{aligned}$$

REMEMBER

$$\mathbf{b} = \tilde{\mathbf{z}} - \mathbf{h}\left(\hat{\mathbf{x}}^{-}, \mathbf{y}\right)$$

$$\mathbf{H}(\hat{\mathbf{x}}^{-}, \mathbf{y}) = \left.\frac{\partial \mathbf{h}(\mathbf{x}, \mathbf{y})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}^{-}}$$

$$\delta\mathbf{x} = \hat{\mathbf{x}}^{+} - \hat{\mathbf{x}}^{-}$$

The residuals are

$$\begin{aligned}\mathbf{v} &\approx \mathbf{H}\delta\mathbf{x} - \mathbf{b} \\ &= \left(\mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}} - \mathbf{I}\right)\mathbf{b}\end{aligned}$$

See Derivation 1 on Moodle

3. Applying Least Squares to Nonlinear Problems

The Linearisation Error

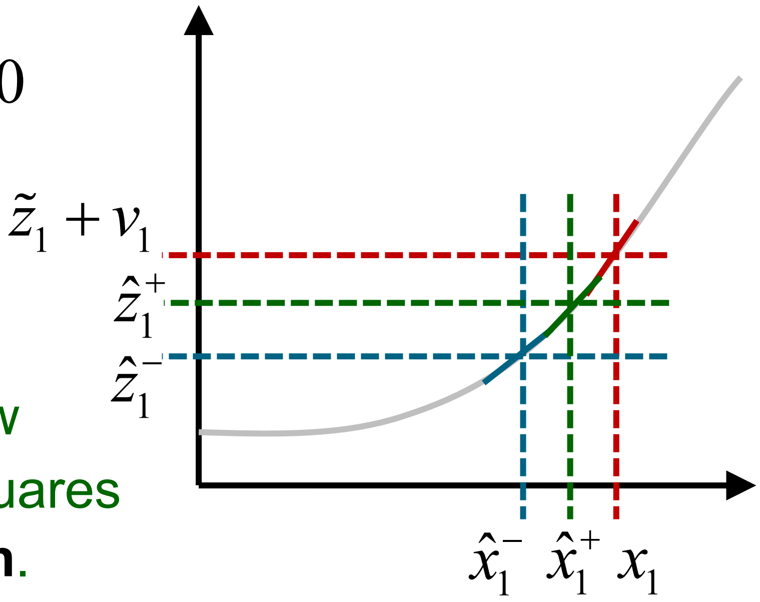
The solution to the nonlinear equation, $\tilde{\mathbf{z}} + \mathbf{v} = \mathbf{h}\left(\hat{\mathbf{x}}^{+}, \mathbf{y}\right)$

is $\hat{\mathbf{x}}^{+} \approx \hat{\mathbf{x}}^{-} + \left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\left(\tilde{\mathbf{z}} - \mathbf{h}\left(\hat{\mathbf{x}}^{-}, \mathbf{y}\right)\right)$

This is only an approximate solution because we have made the assumption

$$\sum_{r=2}^{\infty} \frac{\partial^{r}\mathbf{h}\left(\hat{\mathbf{x}}^{-}, \mathbf{y}\right)}{\partial \mathbf{x}^{r}} \frac{\left[\hat{\mathbf{x}}^{+} - \hat{\mathbf{x}}^{-}\right]^{r}}{r!} \approx 0$$

This is the linearisation approximation



But, $\hat{\mathbf{x}}^{+}$ will be a better solution than $\hat{\mathbf{x}}^{-}$

If we set the predicted states, $\hat{\mathbf{x}}^{-}$, to the new solution, $\hat{\mathbf{x}}^{+}$, and compute another least-squares solution, that will be better. This is **iteration**.

(We must recalculate $\mathbf{H}$)

3. Applying Least Squares to Nonlinear Problems

Iterative Least-Squares (ILS)

Iteration minimises the linearization error for nonlinear least-squares problems

Predict states, $\hat{\mathbf{x}}^-$

Predict measurement, $\hat{\mathbf{z}}^-$

Calculate measurement innovation, $\mathbf{b}$

Calculate measurement matrix, $\mathbf{H}$, using predicted states

Obtain state vector innovation, $\delta\mathbf{x}$, by least-squares

Update state estimates, $\hat{\mathbf{x}}^+ = \hat{\mathbf{x}}^- + \delta\mathbf{x}$

Repeat until convergence ($\delta\mathbf{x}$ within threshold)

$\hat{\mathbf{x}}^- = \hat{\mathbf{x}}^+$

3. Applying Least Squares to Nonlinear Problems

Nonlinear Least-Squares Step-by-Step

Establish: Unknown states (coefficients) to estimate, $\mathbf{x}$

Known parameters, $\mathbf{y}$

Measured parameters $\tilde{\mathbf{z}}$

1) Determine the measurement model: $\mathbf{z} = \mathbf{h}(\mathbf{x}, \mathbf{y})$

2) Predict states, $\hat{\mathbf{x}}^-$

3) Calculate measurement innovation, $\mathbf{b} = \tilde{\mathbf{z}} - \mathbf{h}\left(\hat{\mathbf{x}}^-, \mathbf{y}\right)$

4) Calculate the measurement matrix,

$$\mathbf{H}(\hat{\mathbf{x}}^-, \mathbf{y}) = \left.\frac{\partial \mathbf{h}(\mathbf{x}, \mathbf{y})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}^-}$$

5) Compute the solution, $\hat{\mathbf{x}}^+ \approx \hat{\mathbf{x}}^- + \left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1} \mathbf{H}^{\mathrm{T}}\mathbf{b}$

6) Iterate where necessary

See the Step-by-Step Guide on Moodle

3. Applying Least Squares to Nonlinear Problems

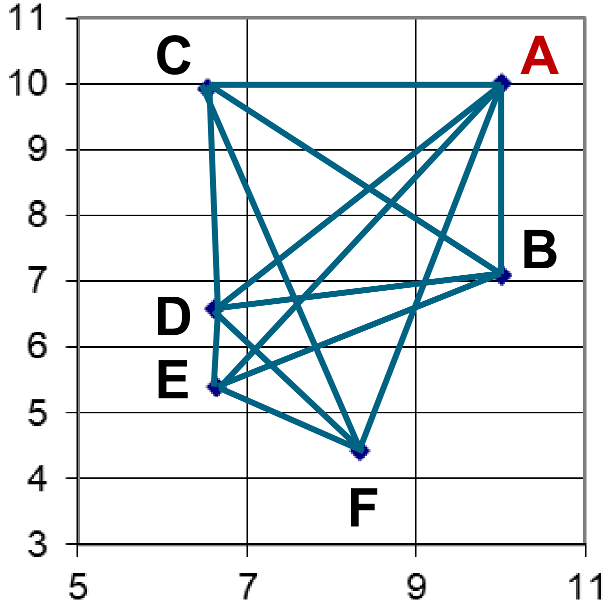
Example 3: Total Station Positioning (1)

A total station measures 13 ranges between 6 points

Coordinates of point A are known

Coordinates of the other 5 points are to be determined

The bearing of A to B (with respect to north) is also measured



States to Estimate, $\mathbf{x}$: E & N coordinates of B, C, D, E & F (10 parameters)

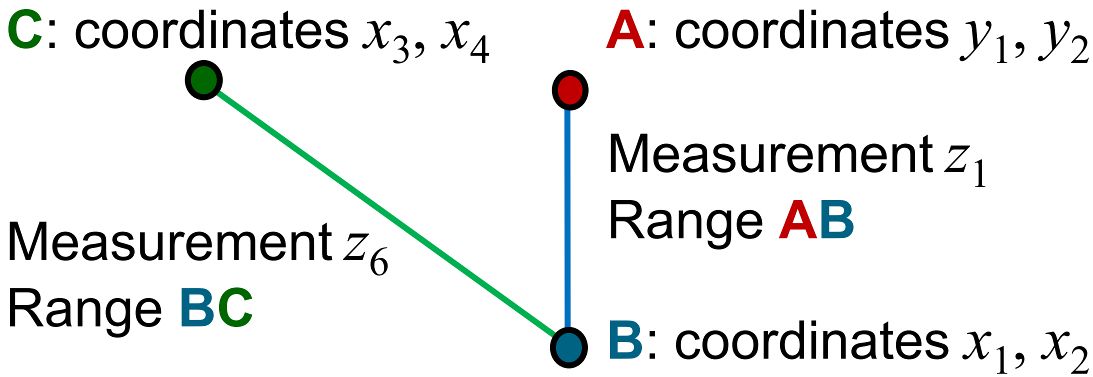
Known Parameters, $\mathbf{y}$: E & N coordinates of A (2 parameters)

Measurements, $\mathbf{z}$: 13 ranges and one bearing

See *RVN Least-Squares Examples.xlsx* on Moodle

Example 3: Total Station Positioning (2)

Step 1: Determine the measurement model - *Ranging*



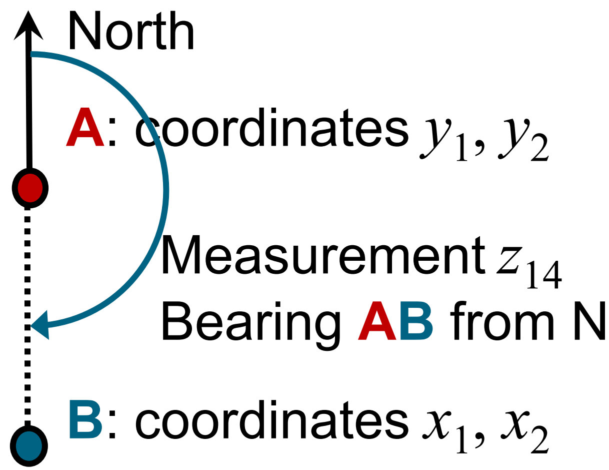
$$\begin{pmatrix} z_1 \\ \vdots \\ z_6 \\ \vdots \end{pmatrix} = \begin{pmatrix} h_1(\mathbf{x},\mathbf{y}) \\ \vdots \\ h_6(\mathbf{x},\mathbf{y}) \\ \vdots \end{pmatrix} = \begin{pmatrix} \sqrt{(x_1-y_1)^2+(x_2-y_2)^2} \\ \vdots \\ \sqrt{(x_3-x_1)^2+(x_4-x_2)^2} \\ \vdots \end{pmatrix}$$

See *RVN Least-Squares Examples.xlsx* on Moodle

3. Applying Least Squares to Nonlinear Problems

Example 3: Total Station Positioning (3)

Step 1: Determine the measurement model - *Bearing*



$$z_{14} = h_{14}(\mathbf{x}, \mathbf{y}) = \arctan_2\left(\left(x_1 - y_1\right), \left(x_2 - y_2\right)\right)$$

Step 2: Predict the states

Point	Easting	Northing
B	10.10	7.10
C	6.50	9.90
D	6.60	6.60
E	6.60	5.40
F	8.30	4.40

Step 3: Calculate the Measurement innovation

Measurement	Measured	Predicted	$\mathbf{b} = \tilde{\mathbf{z}} - \hat{\mathbf{z}}^-$
Range **AB** $= z_1$	2.882	2.902	-0.020
Range **BC** $= z_6$	4.491	4.561	-0.070
Bearing **AB** $= z_{14}$	3.124	3.107	0.017

See *RVN Least-Squares Examples.xlsx* on Moodle

3. Applying Least Squares to Nonlinear Problems

Example 3: Total Station Positioning (4)

Step 4: Calculate the Measurement matrix - *Ranging*

Measurement model: $h_1(\mathbf{x},\mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$

Differentiate with respect to 1st state: $\dfrac{\partial h_1(\mathbf{x})}{\partial x_1} = \dfrac{x_1 - y_1}{\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}}$

See Step-by-Step guide: General Advice for help with the derivatives

Use predicted states for the measurement matrix: $H_{11}(\hat{\mathbf{x}}^-) = \left.\dfrac{\partial h_1(\mathbf{x})}{\partial x_1}\right|_{\mathbf{x}=\hat{\mathbf{x}}^-} = \dfrac{\hat{x}_1^- - y_1}{\sqrt{(\hat{x}_1^- - y_1)^2 + (\hat{x}_2^- - y_2)^2}}$

Simplifying: $H_{11}(\hat{\mathbf{x}}^-) = \dfrac{\hat{x}_1^- - y_1}{\hat{z}_1^-}$ as $\hat{z}_1^- = \sqrt{(\hat{x}_1^- - y_1)^2 + (\hat{x}_2^- - y_2)^2}$

See *RVN Least-Squares Examples.xlsx* on Moodle

3. Applying Least Squares to Nonlinear Problems

Example 3: Total Station Positioning (5)

Step 4: Calculate the Measurement matrix - *Bearing*

Measurement model: $h_{14}(\mathbf{x},\mathbf{y}) = \arctan_2\left((x_1 - y_1),(x_2 - y_2)\right)$

Differentiate with respect to 1st state: $$\frac{\partial h_{14}(\mathbf{x})}{\partial x_1} = \frac{x_2 - y_2}{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

See Step-by-Step guide: General Advice for help with the derivatives

Use predicted states for the measurement matrix: $$H_{14,1}(\hat{\mathbf{x}}^-) = \left.\frac{\partial h_{14}(\mathbf{x})}{\partial x_1}\right|_{\mathbf{x}=\hat{\mathbf{x}}^-} = \frac{\hat{x}_2^- - y_2}{(\hat{x}_1^- - y_1)^2 + (\hat{x}_2^- - y_2)^2}$$

Simplifying: $H_{14,1}(\hat{\mathbf{x}}^-) = \dfrac{\hat{x}_2^- - y_2}{(\hat{z}_1^-)^2}$ as $\hat{z}_1^- = \sqrt{(\hat{x}_1^- - y_1)^2 + (\hat{x}_2^- - y_2)^2}$

See *RVN Least-Squares Examples.xlsx* on Moodle

Example 3: Total Station Positioning (6)

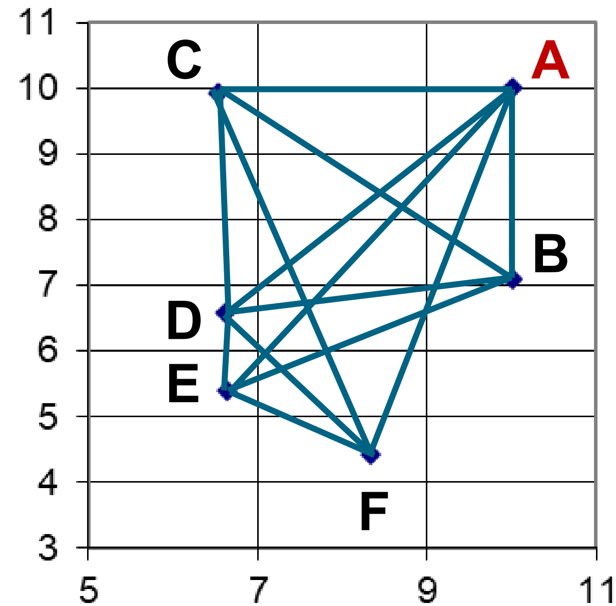
Step 5: Solve

$$\hat{\mathbf{x}}^{+} \approx \hat{\mathbf{x}}^{-} + \left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{b}$$

Step 6: Iterate as needed

After a second iteration…

Point	Easting	Northing
B	10.05	7.11
C	6.52	9.88
D	6.67	6.72
E	6.72	5.36
F	8.43	4.39



See *RVN Least-Squares Examples.xlsx* on Moodle

Contents

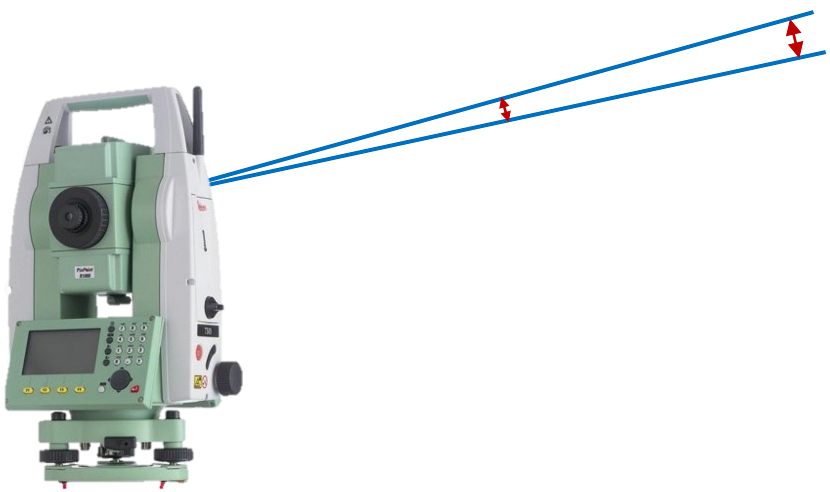
1. Formulating the Problem
2. Linear Least-Squares Estimation
3. Applying Least Squares to Nonlinear Problems
4. Weighted Least-Squares Estimation

4. Weighted Least-Squares Estimation

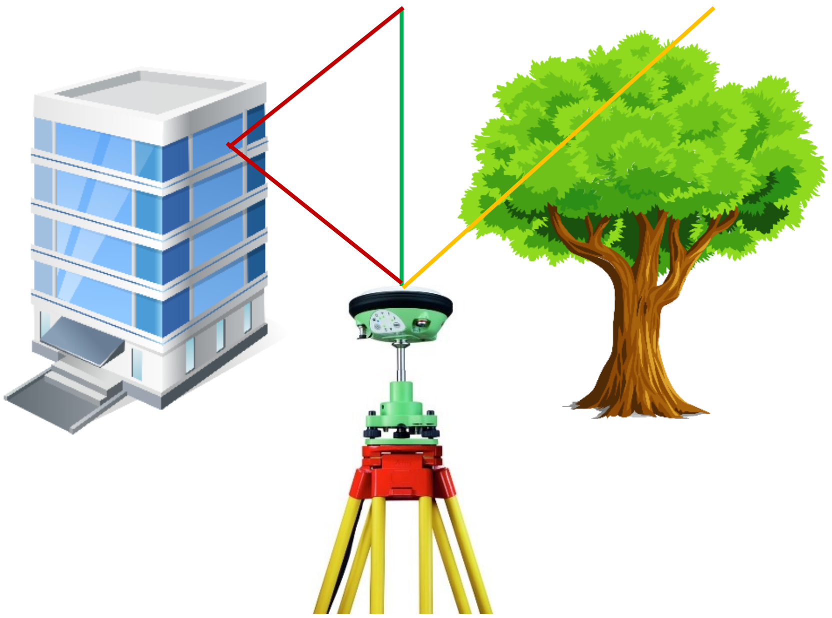
Measurements of Varying Accuracy

Often, some measurements are more precise than others.

Positioning accuracy from angular measurements depends on range



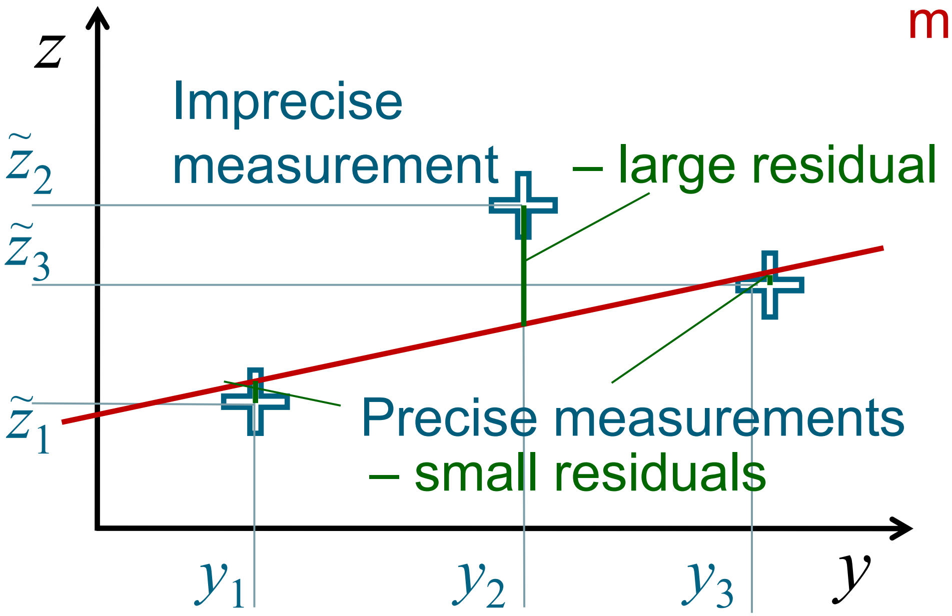
GNSS accuracy can vary between signals



4. Weighted Least-Squares Estimation

Processing Measurements of Varying Accuracy

A simple straight-line example



Equal weighting of measurements is not appropriate where some are much more precise than others.

The function $z = h(\mathbf{x}, y)$ should be closer to the more precise measurements

Residuals should thus be larger for less precise measurements

$\therefore$ *We need to give higher weighting to more precise measurements*

How do we do this?

4. Weighted Least-Squares Estimation

Mean and Variance

The measurement error is given by

Error $\rightarrow$ $\varepsilon = \tilde{z} - z$ $\leftarrow$ True value

Measured value

~ is called 'tilde'

Least-squares estimation assumes measurement errors are zero mean:

Expectation operator
– Gives the mean value of an infinitely large sample

$$\mathrm{E}(\varepsilon) = 0 \qquad \mathrm{E}(\tilde{z}) = z$$

The variance is then: $\sigma_z^2 = \mathrm{E}(\varepsilon^2) = \mathrm{E}\left((\tilde{z} - z)^2\right)$

4. Weighted Least-Squares Estimation

Multiple Measurements

The variances are

$$\sigma_{z1}^2 = \mathrm{E}\left(\varepsilon_1^2\right) = \mathrm{E}\left(\left(\tilde{z}_1 - z_1\right)^2\right)$$
$$\sigma_{z2}^2 = \mathrm{E}\left(\varepsilon_2^2\right) = \mathrm{E}\left(\left(\tilde{z}_2 - z_2\right)^2\right)$$
$$\vdots$$
$$\sigma_{zm}^2 = \mathrm{E}\left(\varepsilon_m^2\right) = \mathrm{E}\left(\left(\tilde{z}_m - z_m\right)^2\right)$$

Different measurements may have different variances or the variances may be the same:

Error sources can affect multiple measurements, so we also need to consider covariance:

$$C_{zij} = \mathrm{E}\left(\left(\tilde{z}_i - z_i\right)\left(\tilde{z}_j - z_j\right)\right) = \boxed{\sigma_{zi}\sigma_{zj}}\rho_{zij}$$

Covariance of i^{th} and j^{th} measurement errors

Measurement error standard deviations

Correlation coefficient

Varies between

–1: fully anticorrelated

0: uncorrelated

+1: fully correlated

4. Weighted Least-Squares Estimation

Measurement Error Covariance Matrix

Expectation of square of the measurement error vector

Comprises variances and covariances of all of the measurement errors

Vector of measured values

$$\mathbf{C_z} = \mathrm{E}\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}\right) = \mathrm{E}\left[\left(\tilde{\mathbf{z}} - \mathbf{z}\right)\left(\tilde{\mathbf{z}} - \mathbf{z}\right)^{\mathrm{T}}\right] = \begin{pmatrix} \sigma_{z1}^2 & C_{z12} & \cdots & C_{zlm} \\ C_{z21} & \sigma_{z2}^2 & \cdots & C_{zlm} \\ \vdots & \vdots & \ddots & \vdots \\ C_{zm1} & C_{zm2} & \cdots & \sigma_{zm}^2 \end{pmatrix}$$

Vector of true values

Off-diagonal elements are covariances

Diagonal elements are variances

Covariance matrices are symmetric: $\mathbf{C_z}^{\mathrm{T}} = \mathbf{C_z}$

This sometimes called the **stochastic model** of the measurements

Introducing Weighted Least-Squares (1)

The **weighted residual** is the ratio of the residual, v, to the measurement error standard deviation, σ_z

The ith weighted residual is v_i / σ_{zi}

$$\sigma_{zi} = \sqrt{\mathrm{E}\left(\varepsilon_i^2\right)} = \sqrt{\mathrm{E}\left[\left(\tilde{z}_i - z_i\right)^2\right]}$$

Where measurement errors are independent…

Minimising the sum of squares of the weighted residuals, not the raw residuals, gives higher weighting to more precise measurements

In general, we minimise $\mathbf{v}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{v}$

$$\mathbf{v}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{v} = \sum_i \frac{v_i^2}{\sigma_{zi}^2} \quad \text{where} \quad \mathbf{C}_{\mathbf{z}} = \begin{pmatrix} \sigma_{z1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{z2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{zm}^2 \end{pmatrix}$$

4. Weighted Least-Squares Estimation

Introducing Weighted Least-Squares (2)

In general, measurement errors are not independent

$$\mathbf{C_z} = \begin{pmatrix} \sigma_{z1}^2 & C_{z12} & \cdots & C_{z1m} \\ C_{z21} & \sigma_{z2}^2 & \cdots & C_{z2m} \\ \vdots & \vdots & \ddots & \vdots \\ C_{zm1} & C_{zm2} & \cdots & \sigma_{zm}^2 \end{pmatrix}$$

← **Stochastic Model**

$$\mathbf{C}_\mathbf{z}^\mathrm{T} = \mathbf{C}_\mathbf{z}$$

By minimising $\mathbf{v}^\mathrm{T}\mathbf{C}_\mathbf{z}^{-1}\mathbf{v}$, errors correlated across different measurements are accounted for

4. Weighted Least-Squares Estimation

Linear Weighted Least-Squares Solution

Derivation is similar to unweighted least-squares

We solve for $\mathbf{x}$ and $\mathbf{v}$: $\tilde{\mathbf{z}} + \mathbf{v} = \mathbf{H}(\mathbf{y})\hat{\mathbf{x}}^{+}$

Constraint: minimise $\mathbf{v}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{v}$

Solution:

$$\boxed{\hat{\mathbf{x}}^{+} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\tilde{\mathbf{z}}}$$

Unweighted solution for comparison

$$\hat{\mathbf{x}}^{+} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\tilde{\mathbf{z}}$$

Residuals:

$$\begin{aligned}\mathbf{v} &= \mathbf{H}\hat{\mathbf{x}}^{+} - \tilde{\mathbf{z}} \\ &= \left(\mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1} - \mathbf{I}\right)\tilde{\mathbf{z}}\end{aligned}$$

See Derivation 2 on Moodle

4. Weighted Least-Squares Estimation

Nonlinear Weighted Least-Squares Solution

The same derivation applies

We solve for $\delta\mathbf{x}$ and $\mathbf{v}$: $\mathbf{b}+\mathbf{v}\approx\mathbf{H}(\hat{\mathbf{x}}^{-},\mathbf{y})\delta\mathbf{x}$

Constraint: minimise $\mathbf{v}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{v}$

Solution:

$$\boxed{\hat{\mathbf{x}}^{+}\approx\hat{\mathbf{x}}^{-}+\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{b}}$$

Where

$$\mathbf{b}=\tilde{\mathbf{z}}-\mathbf{h}\left(\hat{\mathbf{x}}^{-},\mathbf{y}\right)$$

$$\mathbf{H}(\hat{\mathbf{x}}^{-},\mathbf{y})=\left.\frac{\partial\mathbf{h}(\mathbf{x},\mathbf{y})}{\partial\mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}^{-}}$$

$$\delta\mathbf{x}=\hat{\mathbf{x}}^{+}-\hat{\mathbf{x}}^{-}$$

Iterate where necessary

Residuals:

$$\mathbf{v}\approx\mathbf{H}\delta\mathbf{x}-\mathbf{b}$$

$$=\left(\mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}-\mathbf{I}\right)\mathbf{b}$$

Unweighted solution for comparison

$$\hat{\mathbf{x}}^{+}\approx\hat{\mathbf{x}}^{-}+\left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{b}$$

See Derivation 2 on Moodle

4. Weighted Least-Squares Estimation

How Accurate Are the State Estimates?

The state estimation error is given by

Error → $$e = \hat{x}^{+} - x$$ ← True value

↑ Estimated value

^ is called 'caret'

Least-squares estimation assumes state estimation errors are zero mean:

Expectation operator
– Gives the mean value of an infinitely large sample →

$$\mathrm{E}(e) = 0 \qquad \mathrm{E}\left(\hat{x}^{+}\right) = x$$

The variance is then: $\sigma_x^2 = \mathrm{E}\left(e^2\right) = \mathrm{E}\left(\left(\hat{x}^{+} - x\right)^2\right)$

4. Weighted Least-Squares Estimation

Multiple States

The variances are

$$\sigma_{x1}^2 = \mathrm{E}\left(e_1^2\right) = \mathrm{E}\left(\left(\hat{x}_1^+ - x_1\right)^2\right)$$
$$\sigma_{x2}^2 = \mathrm{E}\left(e_2^2\right) = \mathrm{E}\left(\left(\hat{x}_2^+ - x_2\right)^2\right)$$
$$\vdots$$
$$\sigma_{xn}^2 = \mathrm{E}\left(e_n^2\right) = \mathrm{E}\left(\left(\hat{x}_n^+ - x_n\right)^2\right)$$

Different state estimates will usually have different variances

Error sources can affect multiple measurements, so we also need to consider covariance:

$$C_{xij} = \mathrm{E}\left(\left(\hat{x}_i^+ - x_i\right)\left(\hat{x}_j^+ - x_j\right)\right) = \boxed{\sigma_{xi}\sigma_{xj}}\rho_{xij}$$

Covariance of i^{th} and j^{th} state estimation errors

State estimation error standard deviations

Correlation coefficient

Will be non-zero unless the state estimation can be partitioned into separate problems

4. Weighted Least-Squares Estimation

State Estimation Error Covariance Matrix

Expectation of square of the error in the state vector

Comprises variances and covariances of all of the state estimation errors

Vector of estimated values

$$\mathbf{C_x} = \mathrm{E}\left(\mathbf{e}\mathbf{e}^{\mathrm{T}}\right) = \mathrm{E}\left[\left(\hat{\mathbf{x}}^{+} - \mathbf{x}\right)\left(\hat{\mathbf{x}}^{+} - \mathbf{x}\right)^{\mathrm{T}}\right] = \begin{pmatrix} \sigma_{x1}^{2} & C_{x12} & \cdots & C_{x1n} \\ C_{x21} & \sigma_{x2}^{2} & \cdots & C_{x2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{xn1} & C_{xn2} & \cdots & \sigma_{xn}^{2} \end{pmatrix}$$

Vector of true values

Off-diagonal elements are covariances

Diagonal elements are variances

Covariance matrices are symmetric: $\mathbf{C_x}^{\mathrm{T}} = \mathbf{C_x}$

4. Weighted Least-Squares Estimation

State Estimation Error Covariance

Weighted linear least-squares solution: $\hat{\mathbf{x}}^{+} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\tilde{\mathbf{z}}$

Weighted nonlinear solution: $\hat{\mathbf{x}}^{+} \approx \hat{\mathbf{x}}^{-} + \left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{b}$

In both case, we can express the state estimation error, $\mathbf{e}$, as a function of the measurement error, $\boldsymbol{\varepsilon}$, using:

$$\mathbf{e} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\boldsymbol{\varepsilon}$$

The state estimation error covariance is therefore:

$$\mathbf{C}_{\mathbf{x}} = \mathrm{E}\left(\mathbf{e}\mathbf{e}^{\mathrm{T}}\right) = \mathrm{E}\left[\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\right]$$

$$\mathbf{C}_{\mathbf{x}} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathrm{E}\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}\right)\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}$$

$$\mathrm{E}\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}\right) = \mathbf{C}_{\mathbf{z}}$$

$$\mathbf{C}_{\mathbf{x}} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\cancel{\mathbf{C}_{\mathbf{z}}\mathbf{C}_{\mathbf{z}}^{-1}}\mathbf{H}\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}$$

$$\mathbf{C}_{\mathbf{x}} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}\cancel{\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)\left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}}$$

$$\mathbf{C}_{\mathbf{x}} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{H}\right)^{-1}$$

4. Weighted Least-Squares Estimation

Example 4: Total Station Positioning (1)

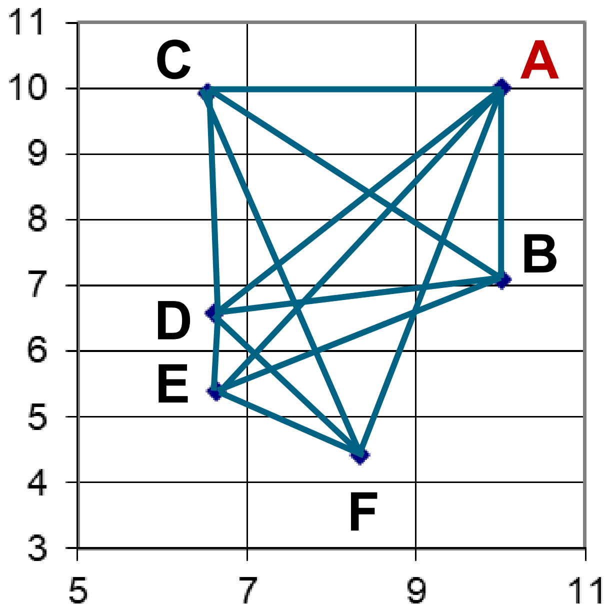
Building on Example 3

A total station measures 13 ranges between 6 points

Coordinates of point A are known

Coordinates of the other 5 points are to be determined

The bearing of A to B (with respect to north) is also measured



States to Estimate, $\mathbf{x}$: E & N coordinates of B, C, D, E & F (10 parameters)

Known Parameters, $\mathbf{y}$: E & N coordinates of A (2 parameters)

Measurements, $\mathbf{z}$: 13 ranges and one bearing

See *RVN Least-Squares Examples.xlsx* on Moodle

4. Weighted Least-Squares Estimation

Example 4: Total Station Positioning (2)

We now have the measurement error standard deviation information:

Ranging measurements: 0.1 m

Bearing measurements: 0.5° = 8.72×10^{-3} rad

All measurements are independent

$$\mathbf{C_z} = \begin{pmatrix} 0.01 & 0 & \cdots & 0 & 0 \\ 0 & 0.01 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0.01 & 0 \\ 0 & 0 & \cdots & 0 & 7.62\times10^{-5} \end{pmatrix}$$

In *RVN Least-Squares Examples.xlsx* on Moodle, a weighted least-squares solution is calculated.

This is the same as the unweighted solution because the bearing measurement is essential for obtaining a unique solution

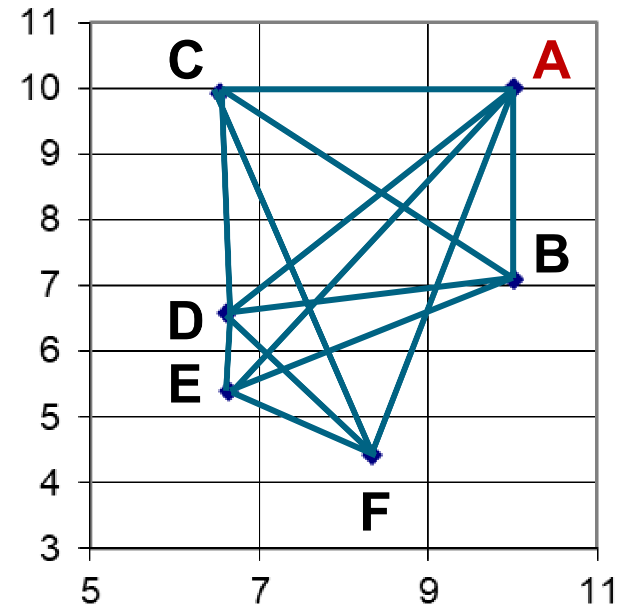
4. Weighted Least-Squares Estimation

Example 4: Total Station Positioning (3)

Calculating the uncertainty of the state estimates using

$$\mathbf{C}_\mathbf{x} = \left(\mathbf{H}^\mathrm{T}\mathbf{C}_\mathbf{z}^{-1}\mathbf{H}\right)^{-1}$$ gives

State	Uncertainty	State	Uncertainty
E_B	0.025	N_B	0.086
E_C	0.090	N_C	0.133
E_D	0.100	N_D	0.130
E_E	0.140	N_E	0.140
E_F	0.197	N_F	0.089



Details in *RVN Least-Squares Examples.xlsx* on Moodle

The east coordinate of B is more accurate than the others due to the angle measurement precision