GPS Parameter Estimation

E. Calais
Purdue University - EAS Department
Civil 3273 – ecalais@purdue.edu

Basic concepts

- Basic problem: We measure range and phase data that are related to the positions of the ground receiver, satellites and other quantities. How do we determine the **"best"** position for the receiver (and other quantities).
- What do we mean by **"best"** estimate?
- Inferring parameters from measurements is an **estimation**
- Two styles of estimation (appropriate for geodetic type measurements):
  - **Parametric estimation** where the quantities to be estimated are the unknown variables in equations that express the observables
  - **Condition estimation** where conditions can be formulated among the observations (*e.g.*, leveling where the sum of the height differences around closed loops must be zero)

Basic concepts

All parametric estimation methods can be broken into a few main steps:

- **Observation equations**: equations that relate the parameters to be estimated to the observed quantities (observables).

  Example: relationship between pseudorange, receiver position, satellite position (implicit in ρ), clocks, atmospheric and ionospheric delays.

- **Stochastic model**: Statistical description that describes the random fluctuations in the measurements (and maybe the parameters).

  Example: covariance matrix that describes the data errors (variance and correlations)

- **Inversion** that solves for the parameters values from the mathematical model consistent with the statistical model.

Observation model

- Observation model are equations relating observables to parameters of model:
  - Observable = function (parameters)
  - Observables should not appear on right-hand-side of equation
- Function is often non-linear, for instance:

$$PR = \sqrt{\left(X_R - X^S\right)^2 + \left(Y_R - Y^S\right)^2 + \left(Z_R - Z^S\right)^2} + \Delta t + \ldots$$

- Then most common method is linearization of function using **Taylor series** expansion.
- Sometimes log linearization for $f=a.b.c$ (products of parameters)

Taylor's series expansion

- Any function y of variables $(x_1, x_2, \ldots, x_n)$ that is indefinitely differentiable an be approximated using a Taylor series expansion as follows:

$$y = f(x_1, x_2, \ldots, x_n)$$

$$y_o + \Delta y = f(X_o) + \frac{\partial f(X_o)}{\partial X} \Delta X + \frac{1}{2!} \frac{\partial^2 f(X_o)}{\partial^2 X} + \ldots \qquad X = (x_1, x_2, \ldots, x_n)$$

observation = data

residual

a priori value

known (gradients of model at X_o)

unknown adjustment

second, third, etc terms (very small)

- The estimation is made using the difference between the observations and the expected values based on **a priori values for the parameters**.
- The estimation returns **adjustments** to a priori parameter values
- Since the linearization is only an approximation, the estimation should be **iterated** until the adjustments to the parameter values are zero.
- For GPS estimation: convergence rate is 100-1000:1 typically (*i.e.*, a 1 meter error in a priori coordinates could result in 1-10 mm of non-linearity error).

Taylor's series expansion

- We wrote: $y_o + \Delta y = f(X_o) + \frac{\partial f(X_o)}{\partial X}\Delta X + \frac{1}{2!}\frac{\partial^2 f(X_o)}{\partial^2 X} + \ldots \qquad X = (x, y, z)$

- For a function of 3 variables x,y,z, neglecting higher order terms:

$$y_o + \Delta y = f(x_o, y_o, z_o) + \frac{\partial f(x_o, y_o, z_o)}{\partial x}\Delta x + \frac{\partial f(x_o, y_o, z_o)}{\partial y}\Delta y + \frac{\partial f(x_o, y_o, z_o)}{\partial z}\Delta z$$

- This can also be written as: $y_o - f(x_o, y_o, z_o) = a_x \Delta x + a_y \Delta y + a_z \Delta z - \Delta y$

- Let's assume we have three observations, then one can write:

$$^1y_o - f(x_o, y_o, z_o) = {}^1a_x \Delta x + {}^1a_y \Delta y + {}^1a_z \Delta z - {}^1\Delta y$$
$$^2y_o - f(x_o, y_o, z_o) = {}^2a_x \Delta x + {}^2a_y \Delta y + {}^2a_z \Delta z - {}^2\Delta y$$
$$^3y_o - f(x_o, y_o, z_o) = {}^3a_x \Delta x + {}^3a_y \Delta y + {}^3a_z \Delta z - {}^3\Delta y$$

Matrix notation

- We wrote:

$$
\begin{aligned}
{}^{1}y_o - f(x_o, y_o, z_o) &= {}^{1}a_x \Delta x + {}^{1}a_y \Delta y + {}^{1}a_z \Delta z - {}^{1}\Delta y \\
{}^{2}y_o - f(x_o, y_o, z_o) &= {}^{2}a_x \Delta x + {}^{2}a_y \Delta y + {}^{2}a_z \Delta z - {}^{2}\Delta y \\
{}^{3}y_o - f(x_o, y_o, z_o) &= {}^{3}a_x \Delta x + {}^{3}a_y \Delta y + {}^{3}a_z \Delta z - {}^{3}\Delta y
\end{aligned}
$$

- This can be expressed in matrix notation as: $L = AX + v$

- With:

$$
L = \begin{bmatrix} {}^{1}y_o - f(x_o, y_o, z_o) \\ {}^{2}y_o - f(x_o, y_o, z_o) \\ {}^{3}y_o - f(x_o, y_o, z_o) \end{bmatrix} \qquad
A = \begin{bmatrix} {}^{1}a_x & {}^{1}a_y & {}^{1}a_z \\ {}^{2}a_x & {}^{2}a_y & {}^{2}a_z \\ {}^{3}a_x & {}^{3}a_y & {}^{3}a_z \end{bmatrix} \qquad
X = \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} \qquad
v = \begin{bmatrix} {}^{1}\Delta y \\ {}^{2}\Delta y \\ {}^{3}\Delta y \end{bmatrix}
$$

Normal equations

- The estimated residuals can be written as: $\hat{v} = L - A\hat{x}$
- We want the solution that minimizes the sum of the squares of the residuals (minimize "L2 norm") => minimize the following functional:

$$J(x) = \sum_{i=i}^{n} v_i^2 = v^T v = \left(L - Ax\right)^T \left(L - Ax\right)$$

- Let us write the derivative of $J(x)$ and set it to zero:

$$\delta J(\hat{x}) = 0$$

$$\delta\left\{\left(L - A\hat{x}\right)^T \left(L - A\hat{x}\right)\right\} = 0$$

$$\delta\left(L - A\hat{x}\right)^T \left(L - A\hat{x}\right) + \left(L - A\hat{x}\right)^T \delta\left(L - A\hat{x}\right) = 0$$

$$\left(-A\delta x\right)^T \left(L - A\hat{x}\right) + \left(L - A\hat{x}\right)^T \left(-A\delta x\right) = 0$$

$$\left(-2A\delta x\right)^T \left(L - A\hat{x}\right) = 0$$

$$\left(\delta x^T A^T\right)\left(L - A\hat{x}\right) = 0$$

$$\delta x^T \left(A^T L - A^T A\hat{x}\right) = 0$$

$$A^T A\hat{x} = A^T L$$

= system of "normal equations"

Least squares solution

- From the normal equations: $A^T A\hat{x} = A^T L$

- The solution for x is: $\hat{x} = \left(A^T A\right)^{-1} A^T L$
  - L = vector of observations
  - A = linear matrix relating parameters to observables (also called design matrix, model matrix, kernel matrix)
  - x = vector of parameters to be estimated
  - v = vector of residuals

- This assumes that the inverse of A^TA exists.

Weighted least-squares

- If the data had no errors, then estimate of x would be perfect. In reality, data have errors and they will map directly into errors on the estimates of x.
- If we know the error in the data, we can form the covariance matrix associated with L:

$$\Sigma_L = \begin{pmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \sigma_n^2 \end{pmatrix}$$

  - Diagonal terms = variance = square of standard deviation
  - Off-diagonal terms = correlations between observation (zero means no correlation)

- Let us define a weight matrix P is : $P = \frac{1}{\sigma_o^2}\Sigma_L^{-1}$
  - $\sigma_o{}^2$ = a priori variance
  - Σ_L = covariance matrix of the observations.

- The least-squares solution now consists of minimizing $v^T P v$ = residuals with larger elements in P are given more weight.

- One can show that the least squares solution is then: $X = (A^T PA)^{-1} A^T PL$

Weighted least-squares

- The law of covariance propagation gives the covariance matrix of the unknowns Σ_X:

$$\Sigma_X = (A^T \Sigma_L^{-1} A)^{-1}$$

- Σ_x has the form:

$$\Sigma_X = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & & \\ \dots & & \ddots & \\ \sigma_{n1} & \dots & & \sigma_n^2 \end{bmatrix}$$

  – Square matrix: size of X times size of X
  – Diagonal terms = variance = square of standard deviation
  – Off-diagonal terms = covariance = degree of correlation between estimated parameters. E.g., if σ_{11} is negative, then a positive error in parameter 1 will be accompanied by a negative error in parameter 2 (and vice versa).

Application to GPS observables

- Pseudorange measurements ${}^{j}R_i(t)$ are modeled as:

$$ {}^{j}R_i(t) = {}^{j}\rho_i(t) + c\left({}^{j}\delta(t) - \delta_i(t)\right) + \Delta I(t) + \Delta T(t) + MP(t) + \varepsilon $$

t = time of epoch
${}^{j}R_i$ = pseudorange measurement
${}^{j}\rho_i$ = satellite-receiver geometric distance
c = speed of light
${}^{j}\delta$ = satellite clock bias

δ_i = receiver clock bias
ΔI = ionospheric propagation error
ΔT = tropospheric propagation error
MP = multipath
ε = receiver noise
(ranges in meters, time in seconds)

- Neglecting propagation, multipath, and receiver errors, eq.(1) becomes:

$$ {}^{j}R_i(t) = {}^{j}\rho_i(t) + c({}^{j}\delta(t) - \delta_i(t)) $$

- The geometric distance between satellite j and receiver i is given by:

$$ {}^{j}\rho_i(t) = \sqrt{\left({}^{j}X(t) - X_i\right)^2 + \left({}^{j}Y(t) - Y_i\right)^2 + \left({}^{j}Z(t) - Z_i\right)^2} \quad or \quad {}^{j}\rho_i(t) = f\left(X_i, Y_i, Z_i\right) $$

with $[{}^{j}X, {}^{j}Y, {}^{j}Z]$ = satellite position, $[X_i, Y_i, Z_i]$ = receiver position in an ECEF coordinate system.

- We need to solve for $[X_i, Y_i, Z_i, \delta_i]$, assuming that we know $[{}^{j}X, {}^{j}Y, {}^{j}Z, {}^{j}\delta]$. Note that the unknowns $[X_i, Y_i, Z_i]$ are **not** linearly related to the observables.

Linearization of pseudorange model

- We need to linearize: $^{j}\rho_i(t) = \sqrt{\left({}^{j}X(t) - X_i\right)^2 + \left({}^{j}Y(t) - Y_i\right)^2 + \left({}^{j}Z(t) - Z_i\right)^2}$ *or* $^{j}\rho_i(t) = f\left(X_i, Y_i, Z_i\right)$
- Let us expand $f(X_i, Y_i, Z_i)$ using a **Taylor's series** with respect to a known position $[X_o, Y_o, Z_o]$:

$$
\begin{aligned}
& f(X_i, Y_i, Z_i) = f(X_o, Y_o, Z_o) \\
& + \frac{\partial f(X_o, Y_o, Z_o)}{\partial X_o} \Delta X_i + \frac{\partial f(X_o, Y_o, Z_o)}{\partial Y_o} \Delta Y_i + \frac{\partial f(X_o, Y_o, Z_o)}{\partial Z_o} \Delta Z_i \\
& + \frac{1}{2!} \frac{\partial^2 f}{\partial x^2} + \ldots
\end{aligned}
$$

- We intentionally truncate the Taylor's expansion after the linear terms.
- Partial derivatives = model gradients at $[X_o, Y_o, Z_o]$
- $[X_o, Y_o, Z_o]$ = known position, in reality approximate position of the site.
- ΔX_i, ΔY_i, ΔZ_i = adjustments, *i.e.* unknowns.

$$
\begin{aligned}
X_i &= X_o + \Delta X_i \\
Y_i &= Y_o + \Delta Y_i \\
Z_i &= Z_o + \Delta Z_i
\end{aligned}
$$

Computing the partial derivatives

Recall from earlier that: $f(X_o,Y_o,Z_o) = \sqrt{\left({}^jX(t)-X_o\right)^2 + \left({}^jY(t)-Y_o\right)^2 + \left({}^jZ(t)-Z_o\right)^2} = {}^j\rho_o(t)$

Recall the chain rule: $\frac{du^n}{dx} = nu^{n-1}\frac{du}{dx}$

Therefore:

$$\frac{\partial f(X_o,Y_o,Z_o)}{\partial X_o} = \frac{\partial\left[\left({}^jX(t)-X_o\right)^2 + \left({}^jY(t)-Y_o\right)^2 + \left({}^jZ(t)-Z_o\right)^2\right]^{1/2}}{\partial X_o}$$

$$= \frac{\partial u(X_o)^{1/2}}{\partial X_o}$$

$$= \frac{1}{2}u(X_o)^{-1/2}\frac{\partial u(X_o)}{\partial X_o}$$

$$= \frac{1}{2\sqrt{u(X_o)}}\frac{\partial\left[\left({}^jX(t)-X_o\right)^2 + \left({}^jY(t)-Y_o\right)^2 + \left({}^jZ(t)-Z_o\right)^2\right]}{\partial X_o}$$

$$= \frac{1}{2\sqrt{u(X_o)}}\frac{\partial\left({}^jX(t)-X_o\right)^2}{\partial X_o}$$

$$= \frac{1}{2\sqrt{u(X_o)}}2\left({}^jX(t)-X_o\right)\frac{\partial\left({}^jX(t)-X_o\right)}{\partial X_o}$$

$$= \frac{{}^jX(t)-X_o}{\sqrt{u(X_o)}}(-1)$$

$$= -\frac{{}^jX(t)-X_o}{{}^j\rho_o(t)}$$

we use:

$$u(X_0) = \left({}^jX(t)-X_o\right)^2 + \left({}^jY(t)-Y_o\right)^2 + \left({}^jZ(t)-Z_o\right)^2$$

Computing the partial derivatives

- The partial derivatives are:

$$\frac{\partial f(X_o,Y_o,Z_o)}{\partial X_o} = -\frac{{}^jX(t)-X_o}{{}^j\rho_o(t)}$$

$$\frac{\partial f(X_o,Y_o,Z_o)}{\partial Y_o} = -\frac{{}^jY(t)-Y_o}{{}^j\rho_o(t)}$$

$$\frac{\partial f(X_o,Y_o,Z_o)}{\partial Z_o} = -\frac{{}^jZ(t)-Z_o}{{}^j\rho_o(t)}$$

- We can now substitute these partial derivatives into the (truncated) Taylor's series expansion:

$$f(X_i,Y_i,Z_i) = f(X_o,Y_o,Z_o) - \frac{{}^jX(t)-X_o}{{}^j\rho_o(t)}\Delta X_i - \frac{{}^jY(t)-Y_o}{{}^j\rho_o(t)}\Delta Y_i - \frac{{}^jZ(t)-Z_o}{{}^j\rho_o(t)}\Delta Z_i$$

- We now have an equation that is linear with respect to the unknowns ΔX_i, ΔY_i, ΔZ_i.

Final linear model

- Let us go back to our pseudorange measurements ${}^{j}R_i(t)$ and rewrite our model equation:

$$ {}^{j}R_i(t) = {}^{j}\rho_o(t) - \frac{{}^{j}X(t) - X_o}{{}^{j}\rho_o(t)}\Delta X_i - \frac{{}^{j}Y(t) - Y_o}{{}^{j}\rho_o(t)}\Delta Y_i - \frac{{}^{j}Z(t) - Z_o}{{}^{j}\rho_o(t)}\Delta Z_i + c\,{}^{j}\delta(t) - c\delta_i(t) $$

- We can rearrange the above equation by separating the known and unknown terms of each side (recall that the satellite clock correction ${}^{j}\delta(t)$ is provided in the navigation message):

$$ {}^{j}R_i(t) - {}^{j}\rho_o(t) - c\,{}^{j}\delta(t) = -\frac{{}^{j}X(t) - X_o}{{}^{j}\rho_o(t)}\Delta X_i - \frac{{}^{j}Y(t) - Y_o}{{}^{j}\rho_o(t)}\Delta Y_i - \frac{{}^{j}Z(t) - Z_o}{{}^{j}\rho_o(t)}\Delta Z_i - c\delta_i(t) $$

- We can simplify the notation by assigning:

$$ {}^{j}a_{Xi} = -\frac{{}^{j}X(t) - X_o}{{}^{j}\rho_o(t)} \qquad {}^{j}a_{Yi} = -\frac{{}^{j}Y(t) - Y_o}{{}^{j}\rho_o(t)} \qquad {}^{j}a_{Zi} = -\frac{{}^{j}Z(t) - Z_o}{{}^{j}\rho_o(t)} \qquad {}^{j}l = {}^{j}R_i(t) - {}^{j}\rho_o(t) - c\,{}^{j}\delta(t) $$

- With 4 satellites visible simultaneously, one can then write the following 4 equations:

$$ \begin{aligned} {}^{1}l &= {}^{1}a_{Xi}\Delta X_i + {}^{1}a_{Yi}\Delta Y_i + {}^{1}a_{Zi}\Delta Z_i - c\delta_i \\ {}^{2}l &= {}^{2}a_{Xi}\Delta X_i + {}^{2}a_{Yi}\Delta Y_i + {}^{2}a_{Zi}\Delta Z_i - c\delta_i \\ {}^{3}l &= {}^{3}a_{Xi}\Delta X_i + {}^{3}a_{Yi}\Delta Y_i + {}^{3}a_{Zi}\Delta Z_i - c\delta_i \\ {}^{4}l &= {}^{4}a_{Xi}\Delta X_i + {}^{4}a_{Yi}\Delta Y_i + {}^{4}a_{Zi}\Delta Z_i - c\delta_i \end{aligned} $$

Problem in matrix notation

- We had, for 4 satellites visible at the same time:

$$
\begin{aligned}
{}^{1}l &= {}^{1}a_{Xi}\Delta X_i + {}^{1}a_{Yi}\Delta Y_i + {}^{1}a_{Zi}\Delta Z_i - c\delta_i \\
{}^{2}l &= {}^{2}a_{Xi}\Delta X_i + {}^{2}a_{Yi}\Delta Y_i + {}^{2}a_{Zi}\Delta Z_i - c\delta_i \\
{}^{3}l &= {}^{3}a_{Xi}\Delta X_i + {}^{3}a_{Yi}\Delta Y_i + {}^{3}a_{Zi}\Delta Z_i - c\delta_i \\
{}^{4}l &= {}^{4}a_{Xi}\Delta X_i + {}^{4}a_{Yi}\Delta Y_i + {}^{4}a_{Zi}\Delta Z_i - c\delta_i
\end{aligned}
$$

- Let us introduce:

$$
A = \begin{bmatrix} {}^{1}a_{Xi} & {}^{1}a_{Yi} & {}^{1}a_{Zi} & -c \\ {}^{2}a_{Xi} & {}^{2}a_{Yi} & {}^{2}a_{Zi} & -c \\ {}^{3}a_{Xi} & {}^{3}a_{Yi} & {}^{3}a_{Zi} & -c \\ {}^{4}a_{Xi} & {}^{4}a_{Yi} & {}^{4}a_{Zi} & -c \end{bmatrix} \qquad \vec{X} = \begin{bmatrix} \Delta X_i \\ \Delta Y_i \\ \Delta Z_i \\ \delta_i \end{bmatrix} \qquad \vec{L} = \begin{bmatrix} {}^{1}l \\ {}^{2}l \\ {}^{3}l \\ {}^{4}l \end{bmatrix}
$$

  - L = vector of *n* observations. Must have at least 4 elements (i.e. 4 satellites), but in reality will have from 4 to 12 elements depending on the satellite constellation geometry.
  - X = vector of u unknowns. Four elements in our case.
  - A = matrix of linear functions of the unknowns (= design matrix), n rows by u columns.

- We can write our problem in a matrix-vector form:

$$
\vec{L} = A\vec{X} + v
$$

Least squares solution

- The data is associated with a variance-covariance matrix Σ_L
- Our problem in a matrix-vector form is written as: $\vec{L} = A\vec{X} \quad , \quad P$
- The least squares solution therefore is given by: $\vec{X} = (A^T PA)^{-1} A^T P\vec{L}$
- P is the weight matrix, defined by: $P = \frac{1}{\sigma_o^2} \Sigma_L^{-1}$

  σ_o^2 = a priori variance
  Σ_L = covariance matrix of the observations.

- The law of covariance propagation gives the covariance matrix of the unknowns Σ_X:

$$\Sigma_X = (A^T \Sigma_L^{-1} A)^{-1}$$

Least squares solution

- The least square solution provides an estimate of: $\vec{X} = \begin{bmatrix} \Delta X_i \\ \Delta Y_i \\ \Delta Z_i \\ \delta_i \end{bmatrix}$

- Once ΔX_i, ΔY_i, ΔZ_i are found, the antenna coordinates $[X_i, Y_i, Z_i]$ are obtained using:

$$X_i = X_o + \Delta X_i$$
$$Y_i = Y_o + \Delta Y_i$$
$$Z_i = Z_o + \Delta Z_i$$

- The covariance matrix of the unknowns Σ_x is:

$$\Sigma_X = (A^T \Sigma_L^{-1} A)^{-1} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} & \sigma_{xt} \\ \sigma_{yx} & \sigma_y^2 & \sigma_{yz} & \sigma_{yt} \\ \sigma_{zx} & \sigma_{zy} & \sigma_z^2 & \sigma_{zt} \\ \sigma_{tx} & \sigma_{ty} & \sigma_{tz} & \sigma_t^2 \end{bmatrix}$$

From ECEF to topocentric

- We can then transform Σ_x from an ECEF frame to a local topocentric frame using the law of variance propagation (disregarding the time-correlated components of Σ_x):

$$\Sigma_T = R\Sigma_X R^T = \begin{bmatrix} \sigma_n^2 & \sigma_{ne} & \sigma_{nu} \\ \sigma_{en} & \sigma_e^2 & \sigma_{eu} \\ \sigma_{un} & \sigma_{ue} & \sigma_u^2 \end{bmatrix}$$

- where R is the rotation matrix:

$$R = \begin{bmatrix} -\sin\varphi\cos\lambda & -\sin\varphi\sin\lambda & \cos\varphi \\ -\sin\lambda & \cos\lambda & 0 \\ \cos\varphi\cos\lambda & \cos\varphi\sin\lambda & \sin\varphi \end{bmatrix}$$

with φ = geodetic latitude of the site, λ = geodetic longitude of the site.

- The DOP factors (Dilution Of Precision) are given by:

$$VDOP = \sigma_u$$

$$HDOP = \sqrt{\sigma_n^2 + \sigma_e^2}$$

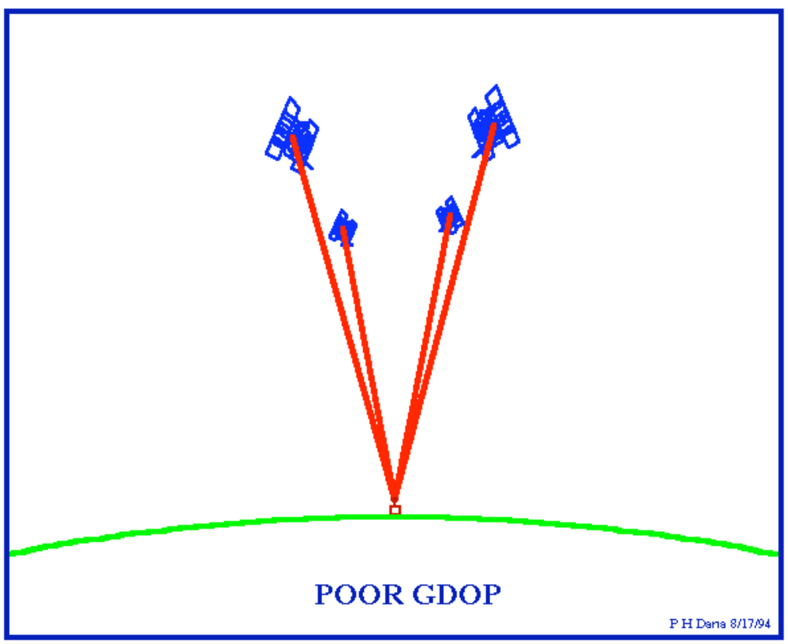
$$PDOP = \sqrt{\sigma_n^2 + \sigma_e^2 + \sigma_u^2} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}$$

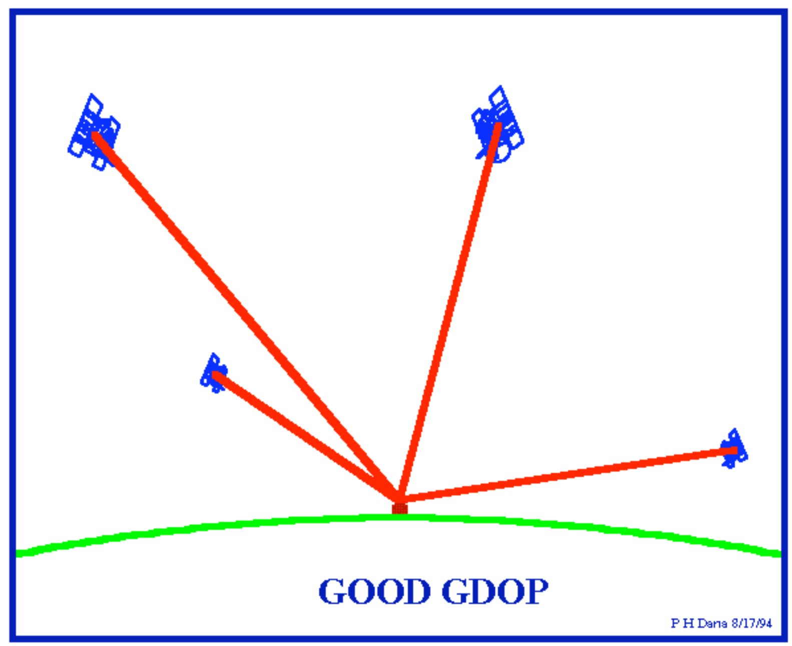
$$TDOP = \sigma_t$$

$$GDOP = \sqrt{\sigma_n^2 + \sigma_e^2 + \sigma_u^2 + \sigma_t^2}$$

Dilution of Precision

– Quantifies the impact of the satellite constellation geometry on position and time:
  - TDOP = time dilution of precision
  - PDOD = position dilution of precision
  - GDOP = geometric dilution of precision (time + position)
– Derived from the diagonal terms of the cofactor matrix ⇒ ~ standard deviations
  - High GDOP ⇒ bad configuration
  - Low GDOP ⇒ good configuration





Positions from phase measurements

- Observable:
  - Remove clock errors ⇒ double difference
  - Dual-frequency receiver ⇒ L_C observable
- Remaining unknowns:
  - ⇒ Antenna position X_i, Y_i, Z_i
  - ⇒ Phase ambiguities: 1 per satellite orbital arc
  - ⇒ Tropospheric delay: 1 zenith total delay parameter every 2 hours, for instance.
- Data:
  - **Static positioning**: the GPS antenna is fixed
    - 1 hour @ 30 sec w/ 8 satellites ⇒ 960 LC observations
    - Unknowns = 12
    - Solve for a system of 960 equations and 12 unknowns (Least squares, Kalman)
    - We can even afford more unknowns, especially if long observation sessions (24 h or more): horizontal tropospheric gradients, orbital parameters, EOP
  - **Kinematic positioning**: the GPS antenna is mobile
    - Data (assuming 8 satellites) = 8 per epoch
    - Unknowns:
      - First epoch = 12
      - As soon as ambiguities are solved = 4 (3)
      - Solving the ambiguities as fast and early as possible is critical
      - Then we can carry them on as we solve for positions.

Error estimation

- Covariance matrix associated with least squares solution:

$$\Sigma_X = (A^T P A)^{-1} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} & \sigma_{xt} \\ \sigma_{yx} & \sigma_y^2 & \sigma_{yz} & \sigma_{yt} \\ \sigma_{zx} & \sigma_{zy} & \sigma_z^2 & \sigma_{zt} \\ \sigma_{tx} & \sigma_{ty} & \sigma_{tz} & \sigma_t^2 \end{bmatrix} \qquad \Sigma_T = R \Sigma_X R^T = \begin{bmatrix} \sigma_n^2 & \sigma_{ne} & \sigma_{nu} \\ \sigma_{en} & \sigma_e^2 & \sigma_{eu} \\ \sigma_{un} & \sigma_{ue} & \sigma_u^2 \end{bmatrix}$$

- Formal errors of the least squares inversion given by:
  - Diagonal terms = variances = (standard deviation)^2
  - Off-diagonal terms = correlations (-1 to 1)

- Interpretation of the formal errors?

Random variables

- Observed and estimated values include random errors = random variables
- Random variables are described by a probability distribution, or probability density = $p(x)$.
- The probability P that a random variable X falls between x and $x+dx$ is found by integrating $p(x)$:

$$P(a \leq X \leq b) = \int_a^b p(x)dx$$

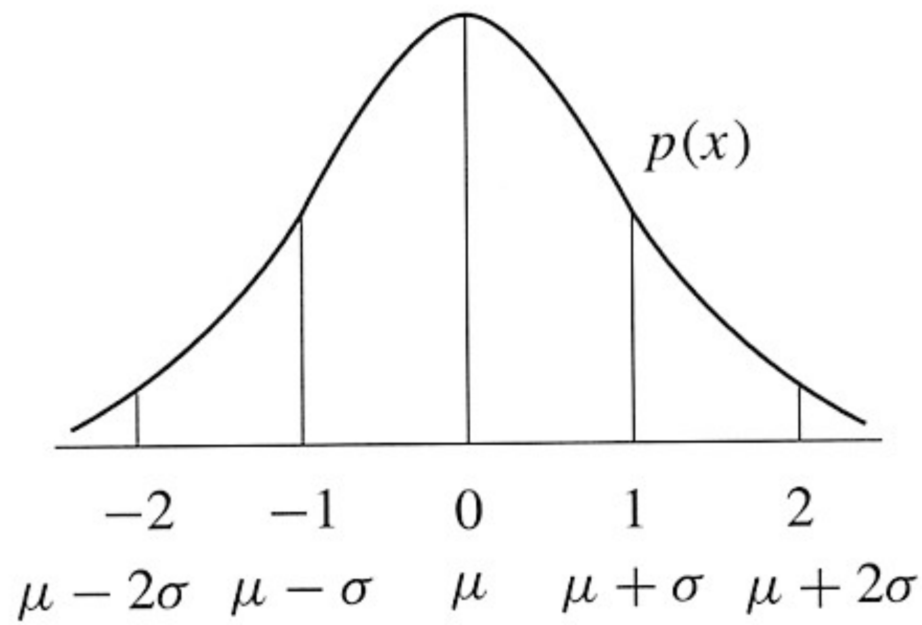
- Of course: $P(-\infty \leq X \leq +\infty) = \int_{-\infty}^{+\infty} p(x)dx = 1$

Normal distribution

- Most important density function = normal (= Gaussian) distribution:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-[x-\mu]^2/2\sigma^2}$$

- Parameters:
  - Mean value μ
  - Standard deviation σ (= measure of scatter around mean)



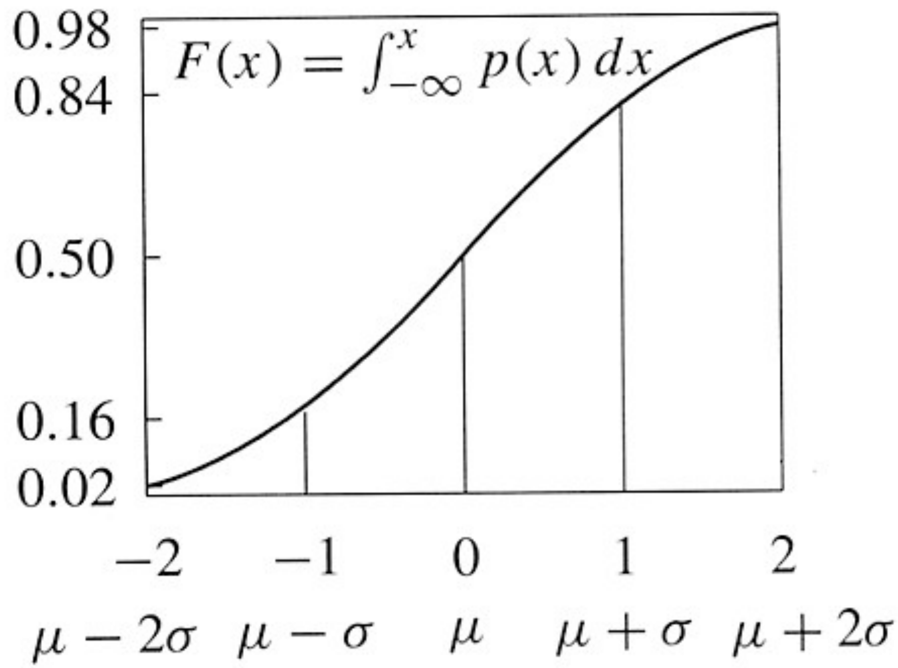
Normal distribution

- Probability that a random sample is below x = cumulative density function $F(x)$:

$$P(X \leq x) = \int_{-\infty}^{x} p(x)dx = F(x)$$

  - At $\mu-\sigma$: $F(x)$ = 0.16
  - At $\mu+\sigma$: $F(x)$ = 0.84
  - Chance of falling between $\mu-\sigma$ and $\mu+\sigma$ = 0.84-0.14 = 0.68 = 68%

- Similarly:
  - 95% corresponds to the chance of falling between $\mu-2\sigma$ and $\mu+2\sigma$
  - 99% corresponds to the chance of falling between $\mu-3\sigma$ and $\mu+3\sigma$



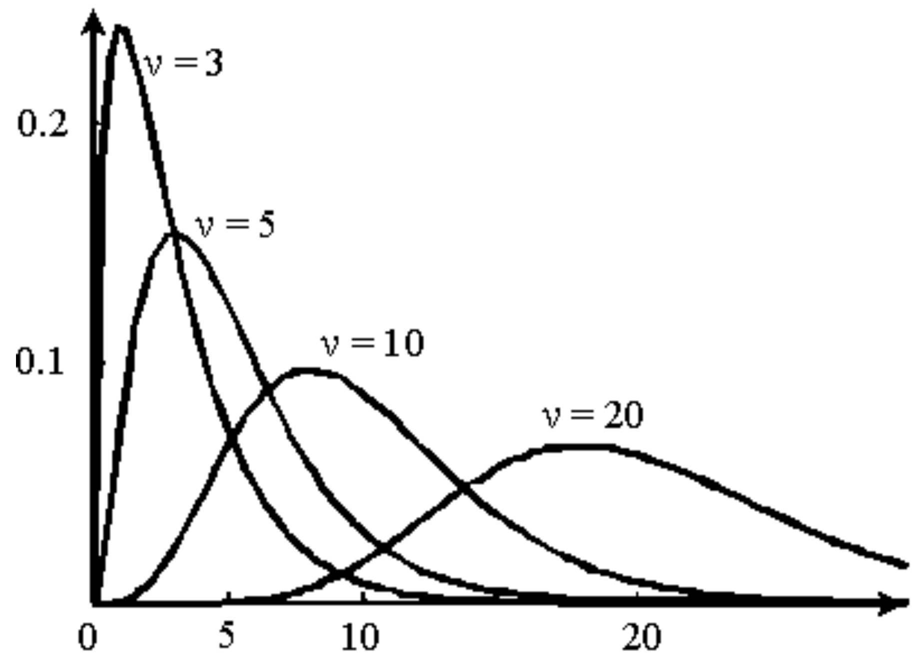
A! valid in 1 dimension…!

Chi-square distribution

- The sum of n independent and normally distributed random variables $x_1^2 + x_2^2 + \ldots + x_n^2$ with zero mean is a random variable (often called χ^2) that follows a chi-square distribution:

$$p_n(x) = \frac{x^{(n/2)-1} e^{-x/2}}{\Gamma(n/2) 2^{n/2}} = \chi_n^2$$

- n = degrees of freedom
- $p(x)$ is no symmetric – approaches normal distribution for $n>30$
- This distribution can be used to calculate the probability K_n that a random variable that follows a chi-square distribution falls within a given interval.



Chi-square distribution

- Assuming independent and normally distributed measurement errors about a zero mean, a least-squares solution (*i.e.*, the fit of N data points y_i $(i=1, ..., N)$ to a model with *M* adjustable parameters a_j $(j=1, ..., M)$) is equivalent to minimizing:

$$\sum_{i=1}^{N}\left[\frac{y_i - y(x_i;a_1,\ldots,a_M)}{\sigma_i}\right]^2$$

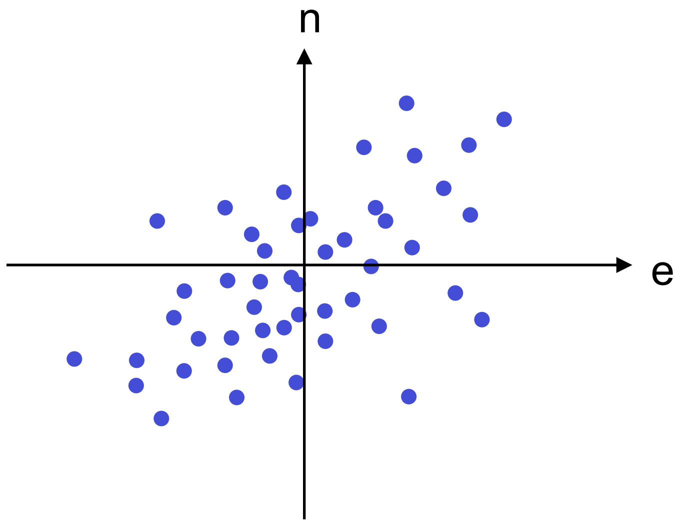
y_i = observations
$y(x_i;a_1,\ldots,a_m)$ = model values
$\Rightarrow y_i - y(x_i;a_1,\ldots,a_m)$ = residuals

(recall that least squares minimizes $v^T P v$)

- Sum of random variables independent and normally distributed about zero $\Rightarrow$ adjustments follows a χ^2 distribution (with $(N\text{-}M)$ degrees of freedom)
- Therefore our position estimates follow a χ^2 distribution – what does that mean for us in practice?

Chi-square distribution

- Example in two-dimensions: solving for horizontal coordinates *(n,e)*
- Blue dots show a series of estimates for the same position (from observations at different times, for instance)
- In two dimensions, the χ^2 distribution can be used to compute the probability that a given estimate falls within a given interval.



c	$K_2(c^2)$
1σ	0.3935
2σ	0.8647
3σ	0.9889

$K_2(c^2)$	c
0.90	2.146
0.95	2.448
0.99	3.035

in 2 dimensions, 1-sigma = 39% confidence…!

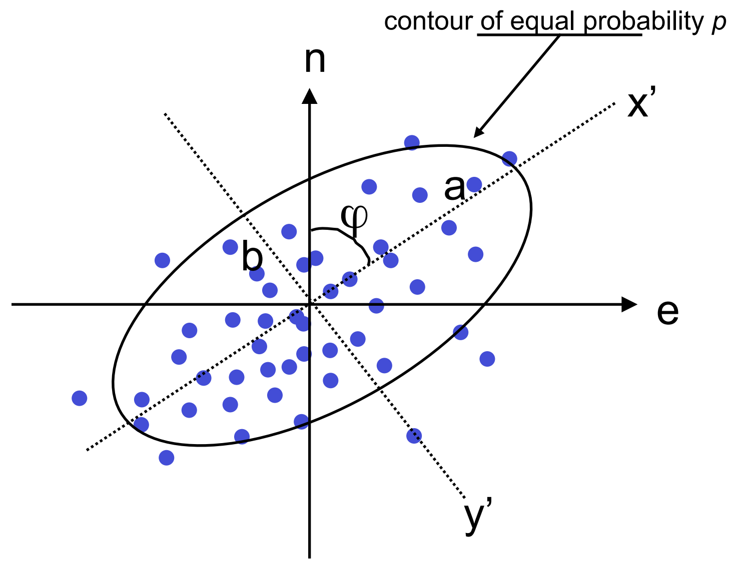
Confidence ellipse

- In 2-dimensions, what is the shape of that interval?
- Geodetic least square problem: estimates *(n,e)* of a particular network point, with the associated covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma_n^2 & \sigma_{ne} \\ \sigma_{en} & \sigma_e^2 \end{bmatrix}$$

- Symmetric matrix => there is a coordinate system with minimum and maximum sigma.
- In that rotated *(x',y')* coordinate system, contours of equal probability to fall in a given interval have the shape of an ellipse:

$$\frac{x'^2}{\left(c\sqrt{\lambda_1}\right)^2} + \frac{y'^2}{\left(c\sqrt{\lambda_2}\right)^2} = 1$$



$$\tan(2\varphi) = \frac{2\sigma_{ne}}{\sigma_n^2 - \sigma_e^2}$$

$$a = c\sqrt{\lambda_1}$$

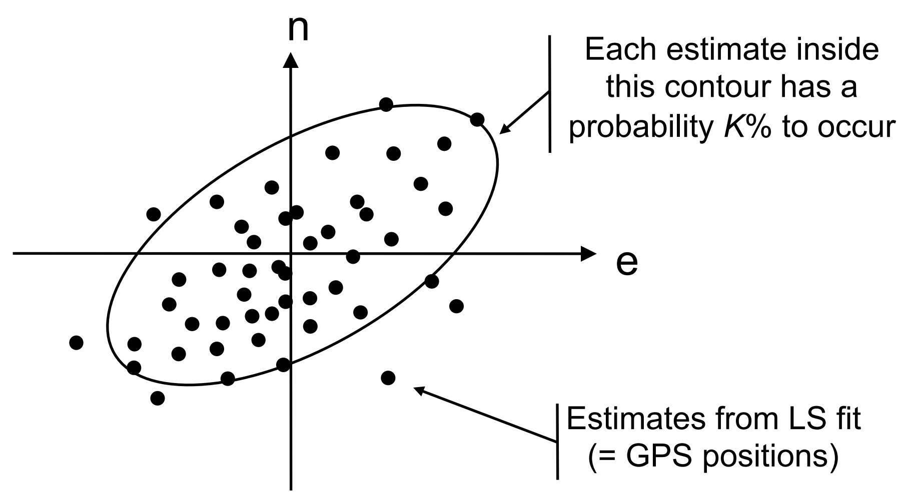
$$b = c\sqrt{\lambda_2}$$

$$\left.\begin{matrix}\lambda_1 \\ \lambda_2\end{matrix}\right\} = \frac{1}{2}\left[\sigma_n^2 + \sigma_e^2 \pm \sqrt{(\sigma_n^2 + \sigma_e^2)^2 - 4\left(\sigma_n^2\sigma_e^2 - \sigma_{ne}^2\right)}\right]$$

Confidence ellipse

- Probability that the estimated point lies within this ellipse? The chi-square cumulative probability K can be used to estimate the probability for the following inequality:

$$\frac{x'^2}{\lambda_1} + \frac{y'^2}{\lambda_2} \leq c^2$$



- Geometrical interpretation of the chi-square: the **confidence ellipse** (or error ellipse)
- Note that there is only 39% chance of being within one-sigma in 2-dimensions
- In geophysics: use 95% confidence = 2.45 sigma

c	$K_2(c^2)$
1σ	0.3935
2σ	0.8647
3σ	0.9889

$K_2(c^2)$	c
0.90	2.146
0.95	2.448
0.99	3.035

in 2 dimensions, 1-sigma = 39% confidence…!

Confidence ellipse

- Covariance matrix:

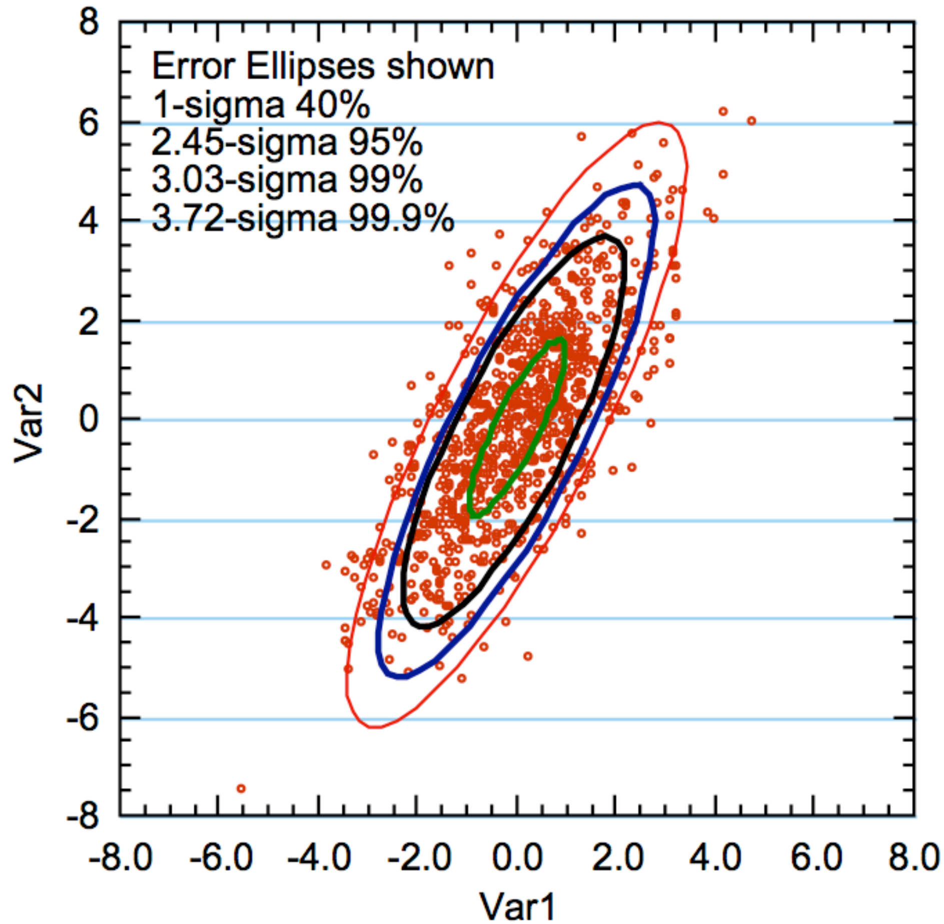
$$\Sigma = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$$

- Eigenvalues:

$$\lambda_1 = 0.87$$
$$\lambda_2 = 3.66$$

- Angle:

$$\varphi = 63$$



From T. Herring

Precision

- Above considerations on errors (="formal" errors) are valid only if the measurement errors are:
  - Independent
  - Normally distributed
- In the case of real (GPS) data:
  - Measurement errors do not necessarily follow a normal distribution…
  - Outliers: data points that are "way off"
    - Least-squares adjustment is still going to try to fit them with a model…
    - Need for careful data editing before inversion (e.g., delete data if error > 3σ)
  - Systematic errors:
    - Do not average out if enough data is taken! ($\neq$ statistical, or random error)
    - Usually very difficult to deal with.
    - E.g.: tribrach calibration, monument deformation.
  - Errors are correlated in time: cf. daily estimates and atmosphere
- Conclusion on formal errors:
  - They are not a realistic representation of the true errors
  - They usually <u>**underestimate**</u> the true error