DATA PROCESSING
(FOR GEOSTATISTICS AND GEOMATICS)

Federica Migliaccio [(1)] – Luigi Mussio [(2)]
[(1)] Politecnico di Milano – DICA – Piazza Leonardo da Vinci, 32 – 20133 Milano
Tel. 02-2399-6507 – Fax 02-2399-6530 – e-mail federica.migliaccio@polimi.it
[(2)] Politecnico di Milano – DICA – Piazza Leonardo da Vinci, 32 – 20133 Milano
Tel. 02-2399-6501 – Fax 02-2399-6602 – e-mail luigi.mussio@polimi.it

Abstract – This paper summarizes some topics about data processing and numerical methods, suitable also for Geostatistic and Geomatics. It starts from elementary statistics, taking into account both the descriptive statistics (in one dimension and bi-dimensional) and statistical inference (i.e. parametric and distribution free tests). Successively it treats estimation theory, considering least squares (in linear and non-linear contexts), numerical control (such conditioning and reliability) and some related problems, such as robust estimators, partitioned models and data sampling. Then it deals with multivariate analysis, starting from cluster analysis, passing through multiple regression and variance analysis, and arriving to sequential tests. After then it deals with interpolation and approximation, starting from the finite element method and spline interpolation, passing through covariance estimation, and arriving to collocation (filtering and prediction), generalized least squares and Kriging. Finally computational statistic and numerical methods are analyzed, with their direct, iterative and sequential methods (particularly for large sparse matrices), special algorithms for regular structures and graph theory. In addition, completely solved exercises are associated to all the principal parts of this paper, supplying a practical support. Furthermore since Data Processing can be placed in a wider frame, characterized by Mathematics and Physics, and they are strictly linked to the Philosophy of Science and the History of Technique, a meeting with Human Sciences concludes this work (while a map collection illustrates a very long path during the history of different peoples and societies).

OVERVIEW

Data processing is historically and theoretically a central part of Survey and Mapping disciplines. Indeed they are characterized by three themes: measurements, models and processing, where the last one is not only an important step, but also the control point of the whole process. In this context, measurements need a control point in term of their accuracy, precision and reliability, while models need a control point in term of optimal design, well conditioning and optimal estimation (or effective robustness). Moreover data processing allows to pass from observed data to model parameters, providing additional information, i.e. metadata, etc. Therefore data processing grew together with Survey and Mapping disciplines, forming an important part of Applied Mathematics.

In particular, the present large spread of Survey and Mapping disciplines requires to be able to adjust networks and image blocks, to reconstruct lines, surfaces and 3D models, to model spatial-temporal events or phenomena, to structure geo-data information and to retrieve and compare them through multilevel, multi-temporal and multi-resolution approaches. For these reasons, a global point of view is strictly necessary and it needs the links with Mathematics and Physics, and a meeting with Human Sciences, where the developments, the research and the applications are clarified. The final goal is to achieve an interdisciplinary perspective, where these specific disciplines offer scientific and technical know-out, and where they receive a political respect from the other disciplines, such as from the whole human society.

PART I – ELEMENTARY STATISTICS

PROBABILITY THEORY

Definition of probability

"Geometric" or Laplace probability: Indistinguishable events $\rightarrow$ Similar probability
(limited to Game Theory)

"Frequency" or Von Mises probability: Limit frequency (empirical)
(true only in probability and therefore contradictory)

"Axioms" or Kolmogorov probability:

1. $0 \le P_A \le 1$
   $A = \varnothing$ and $A = S$ $\Rightarrow$ $P_A = 0$ and $P_A = 1$ respectively
2. $P_A \ge P_B$ if $A \supseteq B$
3. $P_{A+B} = P_A + P_B$ if $AB = \varnothing$
   (mutually exclusive / incompatible events)

Theorems of probability

Sum: $P_{A+B} = P_A + P_B - P_{AB}$ even if $AB \ne \varnothing$
(non-mutually-exclusive/compatible events)

Composed: $P_{AB} = P_A P_B$ only if $\{A, B\}$ independent

$P_{AB} = P_A P_{B|A} = P_{A|B} P_B$ even if $\{A, B\}$ dependent

Bayes: $P_{A|B} = \dfrac{P_{B|A} P_A}{P_B}$ $\qquad$ $P_{B|A} = \dfrac{P_{A|B} P_B}{P_A}$

<u>Note:</u>

- Statistical variables are always results of experiments, therefore concrete, discrete and finite.
- Random variables are models of interpretation and, in general, abstract, unlimited and continuous.
- Exception is given by the random variables of the Game Theory that are abstract, but finite and discrete.

DESCRIPTIVE STATISTICS

SIMPLE STATISTICS VARIABLES

- Moments of k order: $\mu_k = \sum_{i=1}^{n} x_i^k f_i$ $\qquad$ $\mu_k = \dfrac{1}{N}\sum_{i=1}^{n} x_i^k$ if $f_i = \dfrac{1}{N}$
- Central moment of k order: $\bar{\mu}_k = \sum_{i=1}^{n} (x_i - \mu)^k f_i$ $\qquad$ $\bar{\mu}_k = \dfrac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^k$ if $f_i = \dfrac{1}{N}$
- Relation between elementary and central moments:

$$\bar{\mu}_2 = \mu_2 - \mu^2 = \sigma^2 \quad ; \quad \bar{\mu}_2 = \frac{1}{N}\sum_{i=1}^{N} x_i^2 - \mu^2 = \sigma^2$$

$$\bar{\mu}_3 = \mu_3 - 3\mu_2\mu + 2\mu^3 \quad ; \quad \bar{\mu}_3 = \frac{1}{N}\sum_{i=1}^{N} x_i^3 - \frac{3}{N}\mu\sum_{i=1}^{N} x_i^2 + 2\mu^3$$

$$\bar{\mu}_4 = \mu_4 - 4\mu_3\mu + 6\mu_2\mu^2 - 3\mu^4 \quad ; \quad \bar{\mu}_4 = \frac{1}{N}\sum_{i=1}^{N} x_i^4 - \frac{4}{N}\mu\sum_{i=1}^{N} x_i^3 + \frac{6}{N}\mu^2\sum_{i=1}^{N} x_i^2 - 3\mu^4$$

Note: $N =$ number of data;
$n =$ number of the arguments of variable X (number of classes)
$n = 1 + 3.322 \log n$ **Sturges rule** (usually: $5 \le n \le 25$)

Position (or center) indices

The center of order r minimizes the function: $z = \sum_{i=1}^{n} |x_i - c_r|^r f_i$

Mean: $c_2 = \mu = \sum_{i=1}^{n} x_i f_i$ $r = 2$

Median: $c_1 = \mu_e$ $r = 1$ *value that divides the histogram into two equal parts*

Mode: $c_0 = \mu_k$ $r = 0$ *value with the highest frequency*

Dispersion indices

Variance (its square root is called **standard deviation**): $\sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 f_i$ $r = 2$

M.A.V. (mean absolute deviation from median)**:** $\sigma_e = \sum_{i=1}^{n} |x_i - \mu_e| f_i$ $r = 1$

m.a.v. (median of the modules with respect to the median)**:** $\delta_e =$ *median of* $|x_i - \mu_e|$

Amplitude: $\Delta = x_{max} - x_{min}$

Shape Indices

Asymmetry: $\gamma = \dfrac{\sum_{i=1}^{n} (x_i - \mu)^3 f_i}{\sigma^3}$; $\Gamma = \dfrac{c_1 - c_0}{\delta_e}$

$\gamma; \Gamma < 0$ *left tail*
$\gamma; \Gamma = 0$ *symetric*
$\gamma; \Gamma > 0$ *right tail*

Kurtosis: $\beta = \dfrac{\sum_{i=1}^{n} (x_i - \mu)^4 f_i}{\sigma^4}$

$\beta < 3$ *leptokurtic*
$\beta = 3$ *normal*
$\beta > 3$ *platykurtic*

Note: The mean, variance and correlation coefficient (see below) are optimal indices; the median, m.a.v. and Bonferroni indices (again see below) are robust indices.

Chebychev's theorem $P(X) \ge 1 - \dfrac{1}{\lambda^2}$

Normal probability

$\lambda = 1$	$P(X) \ge 0$	$P_N(X) = 0.68 \cong 0.70$
$\lambda = 2$	$P(X) \ge 0.75$	$P_N(X) = 0.95$
$\lambda = 3$	$P(X) \ge 0.89 \cong 0.90$	$P_N(X) = 0.997$
...		
$\lambda = 5$	$P(X) \ge 0.96 \cong 0.95$	
$\lambda = 10$	$P(X) \ge 0.99$	
$\lambda = 18 \cong 20$	$P(X) \ge 0.997$	

Note:

Chebychev's Theorem establishes a comparison between experimental results (i.e. statistical variables)and interpretation models (i.e. random variables), beyond their well known formal identity.

Note of means and mutual variability

Mean Square root : $x_2 = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$

Geometric mean: $x_0 = \sqrt[n]{\prod_{i=1}^{n} x_i}$

Harmonic mean: $x_{-1} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$

Weighted mean: $\bar{x}_p = \frac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i}$

(where: $p_i =$ weight)

Trimmed mean: if $p_i = 0$ for some i

Gini's Delta: $\Delta = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} |x_i - x_j|}{n \cdot (n-1)} \quad (\forall\ j \neq i)$

Concentration ratio:

$$null\ \ 0 \le R = \frac{\Delta}{2\mu} \le 1\ \ max$$

Shannon's Entropy: $H = -\sum_{i=1}^{n} f_i \ log_c\ f_i$

Concentration ratio:

$$null\ \ 0 \le \frac{H}{log_c\ n} \le 1\ \ max$$

$c = 2$ (Information theory)
$c = e$ (continuous random variables)

2D STATISTICAL VARIABLES:

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	y_m	
x_1 i	f_{11}	f_{12}	$\cdots$	f_{1j}	$\cdots$	f_{1m}	p_1
x_2 i	f_{21}	f_{22}	$\cdots$	f_{2j}	$\cdots$	f_{2m}	p_2
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_i	f_{i1}	f_{i2}	$\cdots$	f_{ij}	$\cdots$	f_{im}	p_i
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_n	f_{n1}	f_{n2}	$\cdots$	f_{nj}	$\cdots$	f_{nm}	p_n
	q_1	q_2	$\cdots$	q_j	$\cdots$	q_m	1

Four folder table

where: $f_{ij} =$ relative frequencies $\qquad \sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij} = 1 \qquad N =$ numbers of data

$p_i =$ marginal frequencies X $\qquad \sum_{j=1}^{m} f_{ij} = p_i \qquad m =$ numbers of arguments of variable X

$q_j =$ marginal frequencies Y $\qquad \sum_{i=1}^{n} f_{ij} = q_j \qquad n =$ numbers of arguments of variable Y

Connection (suitable for low dependence)

Contingencies: $c_{ij} = f_{ij} - p_i q_j$

$$-1 \le c_{ij} \le 1$$

Semi contingency mean: $C_0 = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m}\left|c_{ij}\right|$

Bonferroni unilateral indices:

$$\begin{cases} \beta_X = \dfrac{C_0}{1-\sum_{i=1}^{n} p_i^2} & \textit{independence} \quad 0 \le \beta_X \le 1 \quad \textit{perfect dependence} \quad x = h(y) \\ \beta_Y = \dfrac{C_0}{1-\sum_{j=1}^{m} q_j^2} & \textit{independence} \quad 0 \le \beta_Y \le 1 \quad \textit{perfect dependence} \quad y = g(x) \end{cases}$$

Bonferroni bilateral indices:

$$\beta_0 = \sqrt{\beta_X \beta_Y} \qquad \textit{independence} \quad 0 \le \beta_0 \le 1 \quad \textit{perfect bilateral dependence}$$

$$\beta_{-1} = \frac{2\beta_X \beta_Y}{\beta_X + \beta_Y} \qquad \textit{independence} \quad 0 \le \beta_{-1} \le 1 \quad \textit{perfect bilateral dependence}$$

Correlation (linear dependence)

Marginal distribution:

$$X\begin{cases} x_1 & \dots & x_i & \dots & x_n \\ p_1 & \dots & p_i & \dots & p_n \end{cases}$$

$$Y\begin{cases} y_1 & \dots & y_i & \dots & y_n \\ q_1 & \dots & q_i & \dots & q_n \end{cases}$$

Mean of marginal variable X: $\mu_X = \sum_{i=1}^{n} x_i\, p_i$

Mean of marginal variable Y: $\mu_Y = \sum_{j=1}^{n} y_j q_j$

Variance of marginal variable X: $\sigma_X^2 = \sum_{i=1}^{n}\left(x_i - \mu_X\right)^2 p_i$

Variance of marginal variable Y: $\sigma_Y^2 = \sum_{j=1}^{m}\left(y_j - \mu_Y\right)^2 q_j$

Covariance between variables X and Y:

$$\sigma_{XY} = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(x_i - \mu_X\right)\left(y_j - \mu_Y\right) f_{ij} =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j f_{ij} - \mu_X \mu_Y$$

$$\sigma_{XY} = \sum_{k=1}^{N} (x_k - \mu_X)(y_k - \mu_Y) = \sum_{k1}^{N} x_k y_k - \mu_X \mu_Y$$

Linear correlation coefficient:

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \qquad \textit{reverse perfect dependence} \quad -1 \le \rho \le 1 \quad \textit{direct perfect dependence}$$

$$\rho = 0 \quad \textit{linear independence}$$

Regression lines: $Y = aX + b$ $\qquad X = cX + d$

$$a = \frac{\sigma_{XY}}{\sigma_X^2} \qquad c = \frac{\sigma_{XY}}{\sigma_Y^2}$$

$$b = \mu_Y - a\mu_X \qquad d = \mu_X - c\mu_Y$$

Robust regression lines:

$a = median(a_{ij})\ \forall i, j > i$ $\qquad c = median(c_{ij})\ \forall i, j > i$

$b = median(y) - a\ median(x)$ $\qquad d = median(x) - c\ median(y)$

Regression (useful for high-dependencies)

Conditional distributions:

$$X\big|_{y_j} \begin{cases} x_1 & \dots & x_i & \dots & x_n \\ \dfrac{f_{1j}}{q_j} & \dots & \dfrac{f_{ij}}{q_j} & \dots & \dfrac{f_{nj}}{q_j} \end{cases}$$

$$Y\big|_{x_i} \begin{cases} y_1 & \dots & y_i & \dots & y_n \\ \dfrac{f_{i1}}{p_i} & \dots & \dfrac{f_{ij}}{p_i} & \dots & \dfrac{f_{im}}{p_i} \end{cases}$$

Mean of X conditioned by $y = y_j$: $\qquad \mu_{x|_{y_j}} = \sum_{i=1}^{n} \frac{x_i f_{ij}}{q_j} \qquad (i = 1,\dots,n)$

Mean of Y conditioned by $x = x_i$: $\qquad \mu_{y|_{x_i}} = \sum_{j=1}^{m} \frac{y_j f_{ij}}{p_i} \qquad (j = 1,\dots,m)$

Variance of X conditioned by $y = y_j$: $\qquad \sigma^2_{x|_{y_j}} = \sum_{i=1}^{n} \frac{\left(x_i - \mu_{x|_{y_j}}\right)^2 f_{ij}}{q_j} \qquad (i = 1,\dots,n)$

Variance of Y conditioned by $x = x_i$: $\qquad \sigma^2_{y|_{x_i}} = \sum_{j=1}^{m} \frac{\left(y_j - \mu_{y|_{x_i}}\right)^2 f_{ij}}{p_i} \qquad (j = 1,\dots,m)$

Note: The averages of conditioned means coincide with the general averages (but the medians of conditioned medians don't coincide with the general medians, because the median is a procedure and not a linear algorithm).

Variance of conditioned means: $\qquad \sigma^2_{S_{x|y}} = \sum_{j=1}^{m} \left(\mu_{x|_{y_j}} - \mu_X\right)^2 q_j$

Variance of conditioned means: $\sigma^2_{S_{y|x}} = \sum_{i=1}^{n} \left(\mu_{y|x_i} - \mu_Y\right)^2 p_i$

Mean conditioned variances: $\sigma^2_{R_{x|y}} = \sum_{j=1}^{m} \sigma^2_{x|y_j} q_j$

Mean conditioned variances: $\sigma^2_{R_{y|x}} = \sum_{i=1}^{n} \sigma^2_{y|x_i} p_i$

Orthogonal variance decomposition theorem:

$$\sigma^2_{S_{x|y}} + \sigma^2_{R_{x|y}} = \sigma^2_X$$
$$\sigma^2_{S_{y|x}} + \sigma^2_{R_{y|x}} = \sigma^2_Y$$

Pearson unilateral indices (max dependence index: $\sigma^2_R = min \Rightarrow \quad \sigma^2_S = max$):

$$\begin{cases} \eta^2_X = \dfrac{\sigma^2_{S\,X|Y}}{\sigma^2_X} = 1 - \dfrac{\sigma^2_{R\,X|Y}}{\sigma^2_X} & mean\ independence \quad 0 \le \eta^2_x \le 1 \quad perfect\ dependence \quad x = h(y) \\ \eta^2_Y = \dfrac{\sigma^2_{S\,Y|X}}{\sigma^2_Y} = 1 - \dfrac{\sigma^2_{R\,Y|X}}{\sigma^2_Y} & mean\ independence \quad 0 \le \eta^2_Y \le 1 \quad perfect\ dependence \quad y = g(x) \end{cases}$$

Pearson bilateral index (max dependence index):

$$\eta^2 = \frac{\sigma^2_X \eta^2_X + \sigma^2_Y \eta^2_Y}{\sigma^2_X + \sigma^2_Y} \qquad mean\ independence \quad 0 \le \eta^2 \le 1 \quad perfect\ bilateral\ dependence$$

Note: However for the normal distribution, mean independence doesn't imply general independence (as always vice versa guaranteed).

Robust regression

An alternative to the classical regression is the robust one where the median and m.a.v. are used instead of the mean and variance respectively; in this case the variance orthogonal decomposition theorem is, of course, not applicable: however the expected m.a.v. is given by the m.a.v. of the conditioned medians and the residual m.a.v. by the median of the conditioned m.a.v.'s.

TRANSFORMATION OF RANDOM VARIABLES

In general, a transformation of random variable $\underline{X}$ to random variable $\underline{Y}$ can be written as:

$$\underline{Y} = g\left(\underline{X}\right)$$

In particular, a non–singular transformation of random variable $\underline{X}$ ($n\ dimensional$) to random variable $\underline{Y}$ ($m < n\ dimensional$) constitutes a regular transformation (if $m = n$), or otherwise a contraction:

$$\underline{X} = \begin{vmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{vmatrix} \to \underline{Y} = \begin{vmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{vmatrix}$$

Assuming that the mean and covariance matrix of $\underline{X}$ are known:

$$\mu_X = \begin{vmatrix} \mu_{X_1} \\ \mu_{X_2} \\ \vdots \\ \mu_{X_n} \end{vmatrix} \qquad C_{XX} = \begin{vmatrix} \sigma^2_{X_1} & \sigma_{X_1X_2} & \cdots & \sigma_{X_1X_n} \\ \sigma_{X_2X_1} & \sigma^2_{X_2} & \cdots & \sigma_{X_2X_n} \\ \vdots & \vdots & & \vdots \\ \sigma_{X_nX_1} & \sigma_{X_nX_2} & \cdots & \sigma^2_{X_n} \end{vmatrix}$$

one can determine the mean and covariance matrix of $\underline{Y}$.

- Regarding the mean value μ_Y , if $\underline{X}$ is well concentrated around μ_X and in the same zone the function $\underline{Y} = g(\underline{X})$ is gradually varying (i.e. continuous with the first two derivates continuous), it is possible to use the approximate expression:

  $$\mu \cong g(\mu_X)$$

  Note: If $\underline{Y} = g(\underline{X}) = A\underline{X} + b$ is a linear function, the expression $\underline{Y} = g(\mu_X) = A\mu_X + b$ is exact.

- Regarding the covariance matrix C_{YY} , for simplicity, there are two cases:

  1. Linear case: the transformation from X to Y is of linear type: $\underline{Y} = A\underline{X} + b$

     Therefore given coefficient matrix of the linear transformation:

     $$A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix}$$

     the covariance propagation law yields: $C_{YY} = AC_{XX}A^T$

  2. Non-linear case: the transformation from $\underline{X}$ to $\underline{Y}$ is non-linear: $\underline{Y} = F(\underline{X})$

     In this case, it is necessary to introduce the Jacobian matrix, i.e. the matrix of partial derivatives of the functions $g_i(\underline{X})$ with respect to the components of $\underline{X}$ (while the Hessian matrix is the matrix formed by second partial derivates):

     $$J = \begin{vmatrix} \frac{\partial g_1}{\partial X_1} & \frac{\partial g_1}{\partial X_2} & \cdots & \frac{\partial g_1}{\partial X_n} \\ \frac{\partial g_2}{\partial X_1} & \frac{\partial g_2}{\partial X_2} & \cdots & \frac{\partial g_2}{\partial X_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial g_m}{\partial X_1} & \frac{\partial g_m}{\partial X_2} & \cdots & \frac{\partial g_m}{\partial X_n} \end{vmatrix}$$

     and the covariance propagation law yields: $C_{YY} = JC_{XX}J^T$

  Note: According to the variance propagation law and the definition of the variance of moments, the variance of the mean and the variance of the variance of a random variable, not necessarily normal and normal ($\beta = 3$) respectively, in a set of independent data, have the expression:

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \qquad \sigma_{\sigma^2}^2 = \frac{\beta\sigma^4 - \sigma^4}{n} = \frac{2\sigma^4}{n}$$

Note: $m = 1$; $n = 2$

$$\sigma_Y^2 = \left(\frac{\partial g}{\partial X_1}\right)^2 \sigma_{X_1}^2 + \left(\frac{\partial g}{\partial X_2}\right)^2 \sigma_{X_2}^2 + 2\frac{\partial g}{\partial X_1}\frac{\partial g}{\partial X_2}\sigma_{X_1X_2}$$

Examples:

$y = x_1 \pm x_2 \qquad \sigma_y^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 \pm 2\sigma_{x_1x_2}$

$y = x_1x_2 \qquad \sigma_y^2 = \frac{y^2}{x_1^2}\sigma_{x_1}^2 + \frac{y^2}{x_2^2}\sigma_{x_2}^2 + 2\frac{y^2}{x_1x_2}\sigma_{x_1x_2}$

$y = \frac{x_1}{x_2} \qquad \sigma_y^2 = \frac{y^2}{x_1^2}\sigma_{x_1}^2 + \frac{y^2}{x_2^2}\sigma_{x_2}^2 - 2\frac{y^2}{x_1x_2}\sigma_{x_1x_2}$

In both cases, A represents the coefficient matrix of the linear transformation and is the Jacobian, the latter one is the matrix of partial derivatives of functions $g_i(\underline{X})$ with respect to the components of $\underline{X}$, so it's possible to write the following algebraic expressions for the so called variance and covariance propagation law:

$$\sigma_{y_i}^2 = \sum_{k=1}^{n} a_{ik}^2\sigma_{x_k}^2 + 2\sum_{k=1}^{n-1}\sum_{l=k+1}^{n} a_{ik}a_{il}\sigma_{x_kx_l} \quad \forall\, i$$
$$\sigma_{y_iy_j} = \sum_{k=1}^{n} a_{ik}a_{jk}\sigma_{x_k}^2 + \sum_{k=1}^{n-1}\sum_{l=k+1}^{n}\left(a_{ik}a_{jl} + a_{jk}a_{il}\right)\sigma_{x_kx_l} \quad \forall\, i, j \geq i$$

obviously equal to the above matrix expressions.

More generally, from a given random variable $\underline{X}$ with probability density function $p(\underline{X})$, one can obtain the probability density function $p(\underline{Y})$ corresponding to the random variable $\underline{Y}$ by performing a distribution transformation, using the elementary probability conservation theorem; in case the size of the two random variables is identical (i.e. $m = n$), $p(\underline{Y})$ is:

$$p(\underline{Y}) = \frac{p(\underline{X})}{|det\, J|}$$

Note: if: $n = m = 1 \qquad p(y) = \frac{p(x)}{|g'(x)|}$

while in a contraction case where ($m \leq n$) and particularly when $m = 1$ and $n = 2$ the transformation is a sum of the two independent random variables: $y = x + (y - x)$, and has the convolution integral form:

$$p(y) = \int_{-\infty}^{+\infty} p(x)\, p(y - x)\, dx$$

EXAMPLES OF RANDOM VARIABLES

- Normal random variable (for measurement random errors):

One-dimensional probability density function and distribution:

(Standardized variable: $\bar{z} = 0$; $\sigma_z^2 = 1$)

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} \qquad p(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad \text{with} \quad z = \frac{x-\bar{x}}{\sigma}$$

$$P(x) = erf(x) \qquad P(z) = erf(z) = \sigma\, erf(x)$$

Note: Two-dimensional probability density function:

$$p(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\bar{x})^2}{\sigma_x^2} - 2\rho\frac{(x-\bar{x})(y-\bar{y})}{\sigma_x\sigma_y} + \frac{(y-\bar{y})^2}{\sigma_y^2}\right]}$$

semi-axes and orientation of error ellipse:

$$\lambda_{max/\,min} = \frac{\sigma_x^2 + \sigma_y^2}{2} \pm \frac{1}{2}\sqrt{\left(\sigma_x^2 - \sigma_y^2\right)^2 + 4\sigma_{xy}^2} \qquad tan\, 2\vartheta = \frac{-2\sigma_{xy}}{\sigma_x^2 - \sigma_y^2}$$

Note: The characteristic moments of normal random variables only are the mean and the variance; in n–dimensions (with $n > 1$), the only allowed dependence is the linear one and it's expressed by the linear correlation coefficient/s and the variance/s (forming the variance–covariance matrix).

- Uniform random variable (for small segments):

Probability density function:

$$p(x) = \frac{1}{b-a} \qquad a \le x \le b$$

$$p(x) = 0 \qquad x \le a \;\; ; \;\; x \ge b$$

Transformation of normal random variables: $z = erf^{(-1)}(x)$ $\qquad p(z) = N(z)\, p(x)$

- Log - normal random variable (for rare events):

Probability density function:

$$p(x) = \frac{1}{\sqrt{2\pi}ax} e^{-\frac{(ln(x-b))^2}{2a^2}}$$

Transformation of normal random variables: $z = ln\, x$ $\qquad p(z) = x\, p(x)$

Central Limit Theorem and Law of Large Numbers

- The normal distribution, in n-dimension, is characterized by the vectors of means and variances and by the correlation coefficients matrix. The foregoing distribution is symmetric, normal and invariant under linear transformations and convolutions by any infinitesimal perturbations.
- By the Gauss' Theorem or Central Limit Theorem, the normal distribution is the limit distribution of any random variable linear combinations of random variables, under the hypothesis that they are independent with comparable variances (this remark justifies the weighted average of the observations).
- Noting that the Bernoulli's theorem or "law of large numbers" probabilistically guarantees the convergence of statistical variables to random ones, this notion supports the comparison between experimental results (i.e. statistical variables) and interpretative models (i.e. the random variables), beyond their known formal identity.

STATISTICAL INFERENCE – PARAMETRIC TEST

Note: The independence hypothesis is always necessary.

Note: The level of significance α (or error of 1st type) is the risk of rejecting a true hypothesis.

Large sample mean test: **Confidence bounds:**

- Hypothesis H_0: $\mu = \mu_0$ (σ^2 known):

$$\frac{|\overline{x} - \mu_0|}{\frac{\sigma}{\sqrt{N}}} \approx z \qquad \overline{x} - \frac{z\sigma}{\sqrt{N}} \le \mu_0 \le \overline{x} + \frac{z\sigma}{\sqrt{N}}$$

- Hypothesis H_0: $\mu = \mu_0$ (σ^2 unknown):

$$\frac{|\overline{x} - \mu_0|}{\frac{\hat{\sigma}}{\sqrt{N}}} \approx z \qquad \overline{x} - \frac{z\hat{\sigma}}{\sqrt{N}} \le \mu_0 \le \overline{x} + \frac{z\hat{\sigma}}{\sqrt{N}}$$

- Hypothesis H_0: $\mu_X = \mu_Y$ (σ_X^2, σ_Y^2 known):

$$\frac{|\overline{x} - \overline{y}|}{\sqrt{\frac{\sigma_X^2}{N_X} + \frac{\sigma_Y^2}{N_Y}}} \approx z$$

- Hypothesis H_0: $\mu_X = \mu_Y$ (σ_X^2, σ_Y^2 unknown):

$$\frac{|\overline{x} - \overline{y}|}{\sqrt{\frac{\hat{\sigma}_X^2}{N_X} + \frac{\hat{\sigma}_Y^2}{N_Y}}} \approx z$$

Note: In case it is assumed that $\sigma_X^2 = \sigma_Y^2 = \sigma^2$, it is appropriate to give a single estimation of the variance. Hence the expression which is used for the H_0 hypothesis test becomes:

$$\hat{\sigma}^2 = \frac{(N_X - 1)\cdot\hat{\sigma}_X^2 + (N_Y - 1)\cdot\hat{\sigma}_Y^2}{N_X + N_Y - 2} \quad \Rightarrow \quad \frac{|\overline{x} - \overline{y}|}{\hat{\sigma}\sqrt{\frac{1}{N_X} + \frac{1}{N_Y}}} \approx z$$

If H_0 is true, it is appropriate to give a single estimation of the mean as well as of the variance:

$$\overline{\mu} = \frac{N_X \cdot \overline{x} + N_Y \cdot \overline{y}}{N_X + N_Y} \qquad \hat{\sigma}^2 = \frac{(N_X - 1)\cdot\hat{\sigma}_X^2 + (N_Y - 1)\cdot\hat{\sigma}_Y^2 + N_X N_Y(\overline{x} - \overline{y})^2/(N_X + N_Y)}{N_X + N_Y - 1}$$

Power curve of normal distribution

Note: The power of β-test (or error of 2nd type) is the risk of accepting a false hypothesis.

z		0	1	2	2.5	3	4	5
$1-\beta$	$\alpha = 5\%$		15%	50%		85%	$\sim 100\%$	
$1-\beta$	$\alpha = 2\%$		10%	38%		76%	93%	$\sim 100\%$
$1-\beta$	$\alpha = 1\%$		7%	30%	50%	70%	96%	$\sim 100\%$

Normal sample tests: **Confidence bounds:**

- Hypothesis H_0: $\mu = \mu_0$ (σ^2 unknown):

$$\frac{|\bar{x} - \mu_0|}{\frac{\hat{\sigma}}{\sqrt{N}}} \approx t_{N-1} \qquad \bar{x} - \frac{t_\nu \hat{\sigma}}{\sqrt{N}} \le \mu_0 \le \bar{x} + \frac{t_\nu \hat{\sigma}}{\sqrt{N}}$$

- Hypothesis H_0: $\mu_X = \mu_Y$ (σ_X^2, σ_Y^2 unknown):

$$\frac{|\bar{x} - \bar{y}|}{\hat{\sigma}\sqrt{\frac{1}{N_X} + \frac{1}{N_Y}}} = \frac{|\bar{x} - \bar{y}|}{\sqrt{\frac{N_X + N_Y}{N_X N_Y} \cdot \frac{(N_X - 1)\cdot \hat{\sigma}_X^2 + (N_Y - 1)\cdot \hat{\sigma}_Y^2}{N_X + N_Y - 2}}} \approx t_{N_X + N_Y - 2}$$

Note: It is essential that $\sigma_X^2 = \sigma_Y^2 = \sigma^2$, which gives the estimate of $\hat{\sigma}^2$.

- Hypothesis H_0: $\sigma^2 = \sigma_0^2$:

$$\frac{\hat{\sigma}^2}{\sigma_0^2}(N-1) \approx \chi_{N-1}^2 \qquad \frac{\hat{\sigma}^2}{\chi_\nu^{2(+)}} \le \sigma_0^2 \le \frac{\hat{\sigma}^2}{\chi_\nu^{2(-)}}$$

- Hypothesis H_0: $\sigma_X^2 = \sigma_Y^2$:

$$\frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2} \approx F_{N_X - 1, N_Y - 1}$$

- Hhypothesis H_0: $\rho = 0$ (N must be sufficiently large, as proved by the confidence bound chart, drawn by David):

$$\frac{\hat{r}_{XY}}{\sigma_r} = \frac{\hat{r}_{XY}}{\sqrt{\frac{1 - \hat{r}_{XY}^2}{N-2}}} \approx t_{N-2} \qquad \hat{r}_{XY} - t_\nu \sigma_r \le r_0 \le \hat{r}_{XY} + t_\nu \sigma_r$$

Note: For normal samples this is also a test of stochastic independence of X and Y.

- Hypothesis H_0: $\rho = \rho_0$ (under the same conditions):

$$\frac{\left(\frac{1}{2} ln \frac{1 + \hat{r}_{XY}}{1 - \hat{r}_{XY}} - \frac{1}{2} ln \frac{1 + \rho_0}{1 - \rho_0}\right)}{\frac{1}{\sqrt{N-3}}} = \frac{\hat{Z} - \bar{Z}}{\sigma_Z} \approx z$$

DISTRIBUTION – FREE (NON – PARAMETRIC) TESTS

- **Goodness of fit Test:**

Hypothesis H_0: The sample is extracted from a population with known theoretical distribution.

After splitting the sample of N values into m classes, one can compare its absolute frequencies $\hat{F}_i$ with the theoretical probability distributions p_i:

$$\sum_{i=1}^{m} \frac{\left(\hat{F}_i - Np_i\right)^2}{Np_i} = N \sum_{i=1}^{m} \frac{\left(\hat{f}_i - p_i\right)^2}{p_i} \approx \chi^2_{m-h-1}$$

Note: The number h represents the number of nuisance parameters, possibly determined (usually 2). The above comparison is carried out in a similar manner also for a two-dimensional sample (where the number of nuisance parameters is usually 5).

- **Independence Test:**

Hypothesis H_0: Samples X and Y are independent.

Given N pairs of (X, Y) values, ordered in a four fold table and divided into m classes for the X values and into n classes for the Y values, one can compare the absolute frequencies $\hat{F}_{ij}$ with the marginal absolute frequencies $\hat{P}_i \hat{Q}_j$ (where $\nu = nm - (n-1) - (m-1) - 1 = (n-1)(m-1)$):

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left(\hat{F}_{ij} - \frac{\hat{P}_i \hat{Q}_j}{N}\right)^2}{\frac{\hat{P}_i \hat{Q}_j}{N}} = \frac{1}{N} \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left(N\hat{F}_{ij} - \hat{P}_i \hat{Q}_j\right)^2}{\hat{P}_i \hat{Q}_j} =$$

$$= N \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left(\hat{f}_i - \hat{p}_i \hat{q}_j\right)^2}{p_i q_j} \approx \chi^2_{(n-1)(m-1)}$$

- **Test of Kolmogorov-Smirnov:**

Hypothesis H_0: The sample is extracted from a population with known theoretical distribution.

After splitting the sample of N values into m classes, one can compare the sample cumulative relative frequencies $\hat{t}_i$ with the cumulative probabilities P_i given by the theoretical distribution:

$$\max_{i=1,N} \left| \hat{t}_i - P_i \right| \approx D_N$$

Note: $D_{.05} \underset{N>100}{\rightarrow} \frac{1.36}{\sqrt{N}}$ $D_{.01} \underset{N>100}{\rightarrow} \frac{1.63}{\sqrt{N}}$

Note: Strictly speaking, the test should be performed considering one datum at a time, so it is appropriate that any class is as small as possible.

Independence test by the test of Kolmogorov-Smirnov: The above comparison is carried out in a similar manner also for a two-dimensional sample and the same test can be used for the independence test, provided that the products of marginal cumulative frequencies follow the same rules (while the accumulation of two-dimensional frequencies is performed according to the rules of double integrals).

- **Test of Pearson et al. for normality:**

Hypothesis H_0: The sample is extracted from a population with a normal distribution.

The test compares the estimated Skewness and Kurtosis coefficients with the theoretical normal distribution ones ($\bar{\gamma} = 0$ and $\bar{\beta} = 3$):

$$\frac{\hat{\gamma}^2}{6/N} + \frac{(\hat{\beta} - 3)^2}{24/N} \approx \chi_2^2$$

- **Test of Mann-Whitney [1]:**

Hypothesis H_0: $\mu_X = \mu_Y$, comparison of the mean values of two independent variables X and Y.

The data samples are substituted by the corresponding ranks whose values belong to the interval $[1,(N_X + N_Y)]$. The value 1 corresponds to the minimal value and the value $(N_X + N_Y)$ to the maximal one. For a given $\hat{R}_X$ the sum of the ranks of the sample X is:

$$\frac{\hat{R}_X - N_X(N_X + N_Y + 1)/2}{\sqrt{\dfrac{N_X N_Y (N_X + N_Y + 1)}{12}}} \approx z$$

- **Test of Siegel-Tuckey:**

Hypothesis H_0: $\sigma_X^2 = \sigma_Y^2$, comparison of the variances of two independent variables X and Y.

The data samples are substituted by the corresponding ranks whose values belong to the interval $[1,(N_X + N_Y)]$. The value 1 corresponds to the minimal in absolute value residual with respect to the median and the value $(N_X + N_Y)$ to the maximal one. For a given $\hat{R}_X$ the sum of the ranks of the sample X is:

$$\frac{\hat{R}_X - N_X(N_X + N_Y + 1)/2}{\sqrt{\dfrac{N_X N_Y (N_X + N_Y + 1)}{12}}} \approx z$$

- **Test of sign (for central values) [2]:**

Hypothesis H_0: $\mu_X = \mu_Y$.

In the case of the so-called "before and after" studies, i.e. when measuring twice the same sample, one gets two samples, X ("before") and Y ("after") which are not independent. For each pair of values the sign ($+$ or $-$) is determined according to the following rule:

value "before"		value "after"	sign
X	$>$	Y	$-$
X	$<$	Y	$+$
X	$=$	Y	none [3]

Moreover being: $N_p =$ number of "plus" signs, $N_m =$ number of "minus" signs, $N_{tot} = N_p + N_m$, and

[1] This test and the following one are called rank tests.
[2] This test and the following one are defined by Thompson.
[3] The sign vanishes, because the difference is exactly equal to zero.

$\hat{f} = N_p / N_{tot}$, the ratio of the number of "plus" signs over the total number of signs; finally:

$$\frac{\hat{f} - 0.5}{\frac{0.5}{\sqrt{N_{tot}}}} \approx z$$

• Sign test (for the dispersion values):

Hypothesis $H_0 : \quad \sigma_X^2 = \sigma_Y^2$.

Also in this case, the two samples are not independent. The sign ($+$ or $-$) for each pair of the absolute differences between the two samples data and the corresponding medians is determined by the following convention:

$$\begin{array}{cccc} value & & value & sign \\ "before" & & "after" & \\ |X - med_X| & > & |Y - med_Y| & - \\ |X - med_X| & < & |Y - med_Y| & + \\ |X - med_X| & = & |Y - med_Y| & none \end{array}$$ [4]

Moreover being: $N_p =$ number of "plus" signs, $N_m =$ number of "minus" signs, $N_{tot} = N_p + N_m$, and $\hat{f} = N_p / N_{tot}$, the ratio of the number of "plus" signs over the total number of signs; one obtains the same expression as in the Test of sign (for central values).

Note: To compare a relative frequency $\hat{f}$ with a probability $p = p_0$ (or two frequencies: $\hat{f}_1$, $\hat{f}_2$), the following standardizations (of which the of sign tests is a special case) are used:

$$\frac{\hat{f} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{N}}} \approx z \qquad\qquad \frac{\hat{f}_1 - \hat{f}_2}{\sqrt{\frac{\hat{f}_1(1 - \hat{f}_1)}{N_1} + \frac{\hat{f}_2(1 - \hat{f}_2)}{N_2}}} \approx z$$

The same standardization is used to compare the contingency $\hat{c}$ which is given by the difference between the relative frequency $\hat{f}$ and the product of the corresponding relative marginal frequencies: $pq = p_0 q_0$:

$$\frac{\hat{c}}{\sqrt{\frac{p_0 q_0 (1 - p_0 q_0)}{N}}} = \frac{\hat{f} - p_0 q_0}{\sqrt{\frac{p_0 q_0 (1 - p_0 q_0)}{N}}} \approx z$$

- **Test of Spearman [5]:**

Procedure of Spearman's rank correlation coefficients calculation:

- X and Y data sorting

[4] The sign vanishes again, because the difference is exactly equal to zero.
[5] This test is called rank test too.

- assign the ranks separately for each component, in an ascending order
- calculate the differences Δ_i between the ranks of two components
- calculate the Spearman's rank correlation coefficient:

$$\hat{r}_{XY} = 1 - \frac{6}{n(n^2-1)} \sum_{i=1}^{N} \Delta_i^2$$

Hypothesis H_0: $\rho = 0$ (N must be sufficiently large):

$$\frac{\hat{r}_{XY}}{\sqrt{\frac{1-\hat{r}_{XY}^2}{N-2}}} \approx t_{N-2}$$

Hypothesis H_0: $\rho = \rho_0$ (under the same conditions):

$$\frac{\left(\frac{1}{2} ln \frac{1+\hat{r}_{XY}}{1-\hat{r}_{XY}} - \frac{1}{2} ln \frac{1+\rho_0}{1-\rho_0}\right)}{\frac{1}{\sqrt{N-3}}} = \frac{\hat{Z} - \overline{Z}}{\sigma_Z} \approx z$$

- **Welch and Tukey test of mean comparison of normal samples with different variances:**

  - Calculation of expected value: $$\Delta = \frac{|\overline{x} - \overline{y}|}{\sqrt{\frac{\hat{\sigma}_X^2}{N_X} + \frac{\hat{\sigma}_Y^2}{N_Y}}}$$

  - Calculation of the degrees of freedom: $$\nu = \frac{\left(\frac{\hat{\sigma}_X^2}{N_X} + \frac{\hat{\sigma}_Y^2}{N_Y}\right)^2}{\left(\frac{\hat{\sigma}_X^2}{N_X}\right)^2 \frac{1}{N_X+1} + \left(\frac{\hat{\sigma}_Y^2}{N_Y}\right)^2 \frac{1}{N_Y+1}} - 2$$

    Note: If the number of degrees of freedom ν is not an integer number, it is rounded to the nearest one in order to be able to access the usual probability distribution tables.

  - Hypothesis H_0: $\mu_X = \mu_Y$ ($\sigma_X^2, \sigma_Y^2 \neq \sigma_X^2$ unknown):

$$\Delta \approx t_\nu$$

Note: The following asymptotic transformations concerning the normal random variable apply to the χ^2, t of Student and F of Fisher:

$$\sqrt{2\chi_\nu^2} - \sqrt{\nu - 1} \underset{\nu\to\infty}{\rightarrow} z \qquad t_\nu \underset{\nu\to\infty}{\rightarrow} z \qquad \frac{1}{2} ln\, F_{\nu_1\nu_2} \underset{\nu_1\nu_2\to\infty}{\rightarrow} z$$

moreover:

$$F_{\nu_1\nu_2}^{(-)} = \frac{1}{F_{\nu_2\nu_1}^{(+)}}$$

PART II – ESTIMATION THEORY

The estimation theory deals with the adjustment of different models, starting from the fundamental bases of elementary statistics and taking into account some minimum optimal criteria and the variance – covariance propagation law. In particular, polynomial interpolation and finite elements are specific functional models, as well as network structures. These problems involve both linear and non-linear models, as well as optimal or robust criteria; while parallel requirements treat of numerical control (i.e. conditioning and reliability), optimal sampling and simulation / optimization.

Condition equations

Functional model: $By + \Delta = 0$

Stochastic model: $E(y) = \bar{y}$

$D(y) = C_{yy} = \sigma_0^2\, Q_{yy}$ (Gauss–Markov model: $Q_{yy} = Q_{y_0 y_0}$)

Least squares norm:
$$\phi = \frac{1}{2}(\hat{y} - y_0)^T Q_{y_0 y_0}^{-1} (\hat{y} - y_0) + \lambda^T (B\hat{y} + \Delta) =$$
$$= \frac{1}{2}\hat{v}^T P \hat{v} + \lambda^T (B\hat{v} + By_o + \Delta) = min$$

Estimates:
$$\hat{v} = -P^{-1}B^T\left(BP^{-1}B^T\right)^{-1}(By_0 + \Delta) = \hat{y} - y_0$$
$$\hat{y} = y_0 + \hat{v} = y_0 - P^{-1}B^T\left(BP^{-1}B^T\right)^{-1}(By_0 + \Delta)$$

Variance-covariance matrices (based on the covariance propagation law):

$$C_{\hat{v}\hat{v}} = \hat{\sigma}_0^2 P^{-1}B^T\left(BP^{-1}B^T\right)^{-1}BP^{-1} = C_{y_0 y_0} - C_{\hat{y}\hat{y}}$$
$$C_{\hat{y}\hat{y}} = C_{y_0 y_0} - C_{\hat{v}\hat{v}} = \hat{\sigma}_0^2\left(P^{-1} - P^{-1}B^T\left(BP^{-1}B^T\right)^{-1}BP^{-1}\right)$$

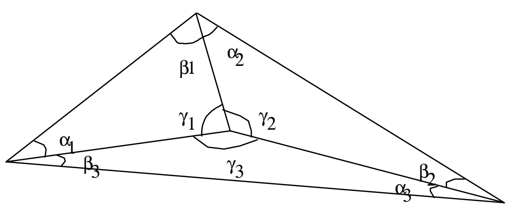
$$C_{\hat{y}\hat{v}} \equiv 0$$

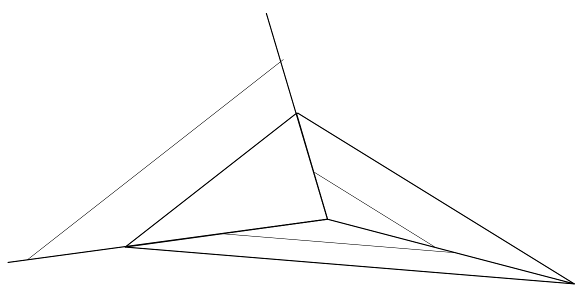
Estimation of sigma-naught: $E\left(v^T P v\right) = TrPE\left(vv^T\right) = \sigma_0^2 TrPQ_{vv} = \sigma_0^2 \nu$

$$\hat{\sigma}_0^2 = \frac{\hat{v}^T P \hat{v}}{\nu}$$

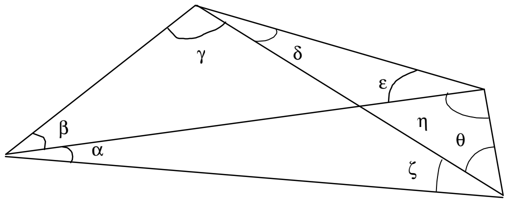
Counter-examples:

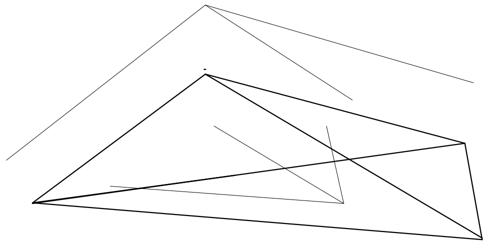
$$sin\alpha_1 sin\beta_1 sin\gamma_1 = sin\alpha_2 sin\beta_2 sin\gamma_2$$



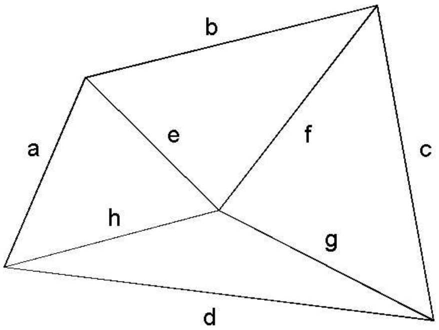


$$\sin a \sin\gamma \sin\varepsilon \sin\vartheta = \sin\beta \sin\delta \sin\eta \sin\zeta$$





$$\sqrt{p(p-a)(p-b)(p-e)} \pm \sqrt{q(q-c)(q-d)(q-e)} =$$
$$= \sqrt{r(r-a)(r-d)(r-f)} + \sqrt{s(s-b)(s-c)(s-f)}$$



where: $p = (a+b+e)/2$; $q = (c+d+e)/2$

$r = (a+d+f)/2$; $s = (b+c+f)/2$

Observation equations (with parameters)

Functional model: $y = Ax + \delta$

Stochastic model: $E(y) = \bar{y}$

$D(y) = C_{yy} = \sigma_0^2 Q_{yy}$ (Gauss–Markov model: $Q_{yy} = Q_{y_0 y_0}$)

Least squares norm:

$$\phi = \frac{1}{2}(\hat{y} - y_0)^T Q_{y_0 y_0}^{-1} (\hat{y} - y_0) + \lambda^T (A\hat{x} + \delta - \hat{v} - y_0) =$$
$$= \frac{1}{2}\hat{v}^T P \hat{v} + \lambda^T (A\hat{x} + \delta - \hat{v} - y_0) = min$$

Estimates:

$$\hat{x} = -(A^T PA)^{-1}(A^T P(\delta - y_0))$$
$$\hat{y} = A\hat{x} + \delta = y_0 + \hat{v}$$
$$\hat{v} = \hat{y} - y_0 = A\hat{x} + \delta - y_0$$

Variance-covariance matrices (based on the covariance propagation law):

$$C_{\hat{x}\hat{x}} = \hat{\sigma}_0^2 (A^T PA)^{-1}$$

$$C_{\hat{y}\hat{y}} = \hat{\sigma}_0^2 A Q_{\hat{x}\hat{x}} A^T = C_{y_0 y_0} - C_{\hat{v}\hat{v}}$$

$$C_{\hat{v}\hat{v}} = C_{y_0 y_0} - C_{\hat{y}\hat{y}} = \hat{\sigma}_0^2 \left(P^{-1} - A Q_{\hat{x}\hat{x}} A^T\right)$$

$$C_{\hat{y}\hat{x}} = -\hat{\sigma}_0^2 A Q_{\hat{x}\hat{x}} = C_{\hat{x}\hat{y}}^T$$
$$C_{\hat{x}\hat{v}} \equiv 0$$
$$C_{\hat{y}\hat{v}} \equiv 0$$

Estimation of sigma-naught: $E\left(v^T P v\right) = TrPE\left(vv^T\right) = \sigma_0^2 TrPQ_{vv} = \sigma_0^2 \nu$

$$\hat{\sigma}_0^2 = \frac{\hat{v}^T P \hat{v}}{m-n}$$

Observation equations (with parameters and constraints)

Functional model: $y = Ax + \delta$

$$Hx + \eta = 0$$

Stochastic model: $E(y) = \bar{y}$

$$D(y) = C_{yy} = \sigma_0^2 Q_{yy}$$ (Gauss–Markov model: $Q_{yy} = Q_{y_0 y_0}$)

Least squares norm: $\phi = \frac{1}{2}\left(\hat{y} - y_0\right)^T Q_{y_0 y_0}^{-1}\left(\hat{y} - y_0\right) + \lambda^T\left(A\hat{x} + \delta - \hat{v} - y_0\right) + \mu^T\left(H\hat{x} + \eta\right) =$

$$= \frac{1}{2}\hat{v}^T P \hat{v} + \lambda^T\left(A\hat{x} + \delta - \hat{v} - y_0\right) + \mu^T\left(H\hat{x} + \eta\right) = min$$

Observation equations (with parameters and over-weighted pseudo-observations)

Functional model: $y = Ax + \delta$

$$z = Hx + \eta = 0$$

Stochastic model: $E(y) = \bar{y} \quad ; \quad E(z) = \bar{z} =$

$$D(y) = C_{yy} = \sigma_0^2 Q_{yy}$$
$$D(z) = C_{zz} = \sigma_0^2 Q_{zz} \cong 0$$ (Gauss–Markov model: $Q_{yy} = Q_{y_0 y_0}$)
$$D(y,z) = 0$$

Least squares norm:

$$\phi = \frac{1}{2}\left(\left(\hat{y} - y_0\right)^T Q_{y_0 y_0}^{-1}\left(\hat{y} - y_0\right) + \left(\hat{z} - z_0\right)^T Q_{z_0 z_0}^{-1}\left(\hat{z} - z_0\right)\right) +$$
$$+ \lambda^T\left(A\hat{x} + \delta - \hat{v} - y_0\right) + \mu^T\left(H\hat{x} + \eta - \hat{u} - z_0\right) =$$
$$= \frac{1}{2}\left(\hat{v}^T P \hat{v} + \hat{u}^T Q \hat{u}\right) + \lambda^T\left(A\hat{x} + \delta - \hat{v} - y_0\right) + \mu^T\left(H\hat{x} + \eta - \hat{u} - z_0\right) = min$$

Estimates:

$$\hat{x} = -\left(A^T P A + H^T Q H\right)^{-1}\left(A^T P\left(\delta - y_0\right) + H^T Q \eta\right)$$

$$\hat{y} = A\hat{x} + \delta = y_0 + \hat{v}$$

$$\hat{v} = \hat{y} - y_0 = A\hat{x} + \delta - y_0$$

$$\hat{z} = H\hat{x} + \eta = \hat{u} = 0$$

Variance-covariance matrices (based on the covariance propagation law):

$$C_{\hat{x}\hat{x}} = \hat{\sigma}_0^2\left(A^T PA + H^T QH\right)^{-1}$$

$$C_{\hat{y}\hat{y}} = \hat{\sigma}_0^2 AQ_{\hat{x}\hat{x}}A^T = C_{y_0y_0} - C_{\hat{v}\hat{v}}$$
$$C_{\hat{v}\hat{v}} = C_{y_0y_0} - C_{\hat{y}\hat{y}} = \hat{\sigma}_0^2\left(P^{-1} - AQ_{\hat{x}\hat{x}}A^T\right)$$

$$C_{\hat{y}\hat{x}} = \hat{\sigma}_0^2 AQ_{\hat{x}\hat{x}} = C^T{}_{\hat{x}\hat{y}}$$
$$C_{\hat{x}\hat{v}} \equiv 0 \quad ; \quad C_{\hat{y}\hat{v}} \equiv 0$$

$$C_{\hat{z}\hat{z}} = C_{z_0z_0} - C_{\hat{u}\hat{u}} = 0$$
$$C_{\hat{u}\hat{u}} = C_{z_0z_0} - C_{\hat{z}\hat{z}} = 0$$

$$C_{\hat{x}\hat{z}} = 0$$
$$C_{\hat{x}\hat{u}} \equiv 0 \quad ; \quad C_{\hat{y}\hat{u}} \equiv 0$$

$$C_{\hat{y}\hat{z}} = 0 \quad ; \quad C_{\hat{v}\hat{z}} \equiv 0$$
$$C_{\hat{v}\hat{u}} \equiv 0 \quad ; \quad C_{\hat{z}\hat{u}} = 0$$

Estimation of sigma-naught: $E\left(v^T Pv\right) = TrPE\left(vv^T\right) = \sigma_0^2 TrPQ_{vv} = \sigma_0^2 \nu$

$$\hat{\sigma}_0^2 = \frac{\hat{v}^T P\hat{v}}{m - n + l}$$

where:

$$A = \begin{bmatrix} A \\ H \end{bmatrix} \quad ; \quad \delta = \begin{bmatrix} \delta \\ \eta \end{bmatrix} \quad ; \quad P = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}$$

$$y_0 = \begin{bmatrix} y_0 \\ z_0 \end{bmatrix} \quad ; \quad \hat{y} = \begin{bmatrix} \hat{y} \\ \hat{z} \end{bmatrix} \quad ; \quad \hat{v} = \begin{bmatrix} \hat{v} \\ \hat{u} \end{bmatrix}$$

Observation equations (with parameters and generic pseudo-observations suitable for additional information and regularizations)

Estimates:

$$\hat{x} = -\left(A^T PA + H^T QH\right)^{-1}\left(A^T P\left(\delta - y_0\right) + H^T Q\eta\right)$$

$$\hat{y} = A\hat{x} + \delta = y_0 + \hat{v}$$

$$\hat{v} = \hat{y} - y_0 = A\hat{x} + \delta - y_0$$

$$\hat{z} = H\hat{x} + \eta = z_0 + \hat{u}$$
$$\hat{u} = \hat{z} - z_0 = H\hat{x} + \eta - z_0$$

Variance-covariance matrices (based on the covariance propagation law):

$$C_{\hat{x}\hat{x}} = \hat{\sigma}_0^2 \left(A^T PA + H^T QH\right)^{-1}$$
$$C_{\hat{y}\hat{y}} = \hat{\sigma}_0^2 AQ_{\hat{x}\hat{x}} A^T = C_{y_0 y_0} - C_{\hat{v}\hat{v}}$$
$$C_{\hat{v}\hat{v}} = C_{y_0 y_0} - C_{\hat{y}\hat{y}} = \hat{\sigma}_0^2 \left(P^{-1} - AQ_{\hat{x}\hat{x}} A^T\right)$$

$$C_{\hat{y}\hat{x}} = \hat{\sigma}_0^2 AQ_{\hat{x}\hat{x}} = C^T{}_{\hat{x}\hat{y}}$$
$$C_{\hat{x}\hat{v}} \equiv 0 \quad ; \quad C_{\hat{y}\hat{v}} \equiv 0 \quad ; \quad C_{\hat{x}\hat{u}} \equiv 0 \quad ; \quad C_{\hat{y}\hat{u}} \equiv 0$$

$$C_{\hat{z}\hat{z}} = \hat{\sigma}_0^2 HQ_{\hat{x}\hat{x}} H^T = C_{z_0 z_0} - C_{\hat{u}\hat{u}}$$
$$C_{\hat{u}\hat{u}} = C_{z_0 z_0} - C_{\hat{z}\hat{z}} = \hat{\sigma}_0^2 \left(Q^{-1} - HQ_{\hat{x}\hat{x}} H^T\right)$$

$$C_{\hat{x}\hat{z}} = -\hat{\sigma}_0^2 Q_{\hat{x}\hat{x}} H^T = C_{\hat{z}\hat{x}}^T$$
$$C_{\hat{y}\hat{z}} = \hat{\sigma}_0^2 AQ_{\hat{x}\hat{x}} H^T = C_{\hat{z}\hat{y}}^T$$
$$C_{\hat{v}\hat{u}} = -C_{\hat{y}\hat{z}} \qquad ; \quad C_{\hat{v}\hat{z}} = 0 \quad ; \quad C_{\hat{z}\hat{u}} = 0$$

Estimation of sigma-naught: $\hat{\sigma}_0^2 = \dfrac{\hat{v}^T P\hat{v} + \hat{u}^T Q\hat{u}}{m - n + l}$

Complements

- Direct constraints and pseudo-observations:

$$\Pi\,\hat{x} = \Pi\left(x_0 + \hat{u}\right)$$
$$C_{x_0 x_0} = \hat{\sigma}_0^2\, \Pi\, Q^{-1} \Pi$$

$$\hat{x} = -\left(A^T PA + \Pi Q \Pi\right)^{-1} \left(A^T P\left(\delta - y_0\right) + \Pi Q\, x_0\right)$$
$$C_{\hat{x}\hat{x}} = \hat{\sigma}_0^2 \left(A^T PA + \Pi Q \Pi\right)^{-1}$$

- Slack parameters:

$$\hat{y} = y_0 + \hat{v} = A\hat{x}_1 + B\hat{x}_2 + \delta$$

$$\hat{y} = y_0 + \hat{v} = A\hat{x}_1 + B\hat{x}_2 + \delta$$
$$\hat{z} = z_0 + \hat{u} = H\hat{x}_1 + K\hat{x}_2 + \eta$$

$$\hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} \quad ; \quad A = \begin{bmatrix} A & B \end{bmatrix}$$

$$\hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} \quad ; \quad A = \begin{bmatrix} A & B \end{bmatrix} \quad ; \quad H = \begin{bmatrix} H & K \end{bmatrix}$$

- Supplementary observations:

$$\hat{y} = y_0 + \hat{v} = A\hat{x}_1 + Bz_0 + \delta \qquad A\hat{x}_1 + By_0 + \delta - \hat{v} = 0$$

$$\hat{y} = y_0 + \hat{v} = A\hat{x}_1 + B\hat{x}_2 + \delta \qquad\qquad A\hat{x}_1 + B\hat{x}_2 + \delta = 0$$

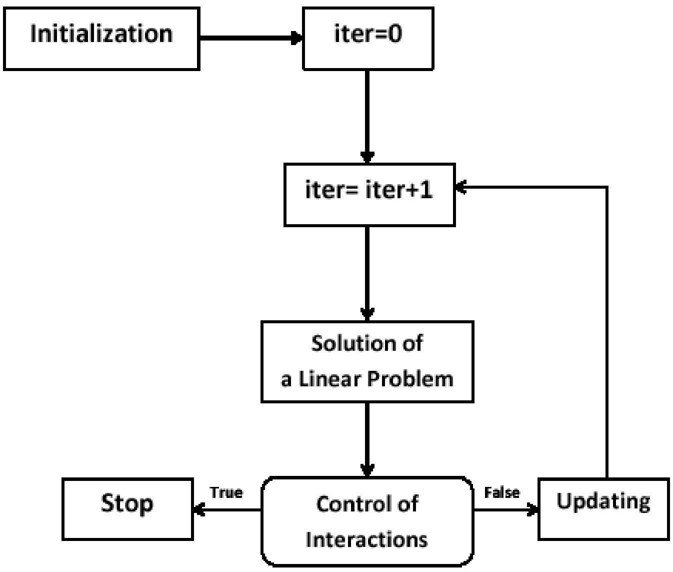
$$\hat{z} = z_0 + \hat{u} = \hat{x}_2 \qquad\qquad \hat{y} = y_0 + \hat{u} = \hat{x}_2$$

$$P = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} \quad ; \quad P = \begin{bmatrix} Q & 0 \\ 0 & P \end{bmatrix}$$

then according to slack parameter technicalities:

$$\hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} \quad ; \quad A = \begin{bmatrix} A & B \\ 0 & I \end{bmatrix} \quad ; \quad \delta = \begin{bmatrix} \delta \\ 0 \end{bmatrix} \qquad P = \begin{bmatrix} Q & 0 \\ 0 & P \end{bmatrix}$$

Non-linear problems



- Functional model: Newton-Fourier method

$$y = F(x)$$

$$y = F(\tilde{x}) + F_x(\tilde{x})(x - \tilde{x}) = \tilde{y} + J(\tilde{x})(x - \tilde{x})$$

$$\max_{i=1,n} = \left(|\hat{x} - \tilde{x}|\right)_i$$

Tikhonov regularization

$$\left(A^T PA + \lambda\, I\right)x + A^T P\left(\delta - y^0\right) = 0$$

Levenberg-Marquardt algorithm
(suitable to accelerate the convergence in non-linear problems)

$$\left(A^T PA + \lambda\, diag\left(A^T PA\right)\right)x + A^T P\left(\delta - y^0\right) = 0$$

- Stochastic model (Gauss-Helmert model for weight reproduction):

$$C_{yy} = \sigma_0^2\left(I \otimes \sigma_i^2 P_i^{-1} / \sigma_0^2\right)$$

$$\hat{\sigma}_i^2 = \frac{(\Pi_i\hat{v})^T \Pi_i \left(I \otimes \sigma_0^2 \frac{P_j}{\sigma_j^2} \right) \Pi_i \; (\Pi_i\hat{v})}{\nu_i} \qquad\qquad i = 1,h$$

$$\nu_i\sigma_i^2 = E\left((\Pi_i\hat{v})^T \Pi_i \left(I \otimes \sigma_0^2 \frac{P_j}{\sigma_j^2} \right) \Pi_i (\Pi_i\hat{v}) \right) = TrE\left((\Pi_i\hat{v})^T \Pi_i \left(I \otimes \sigma_0^2 \frac{P_j}{\sigma_j^2} \right) \Pi_i (\Pi_i\hat{v}) \right)$$

$$= Tr\left(\Pi_i \left(I \otimes \sigma_0^2 \frac{P_j}{\sigma_j^2} \right) \Pi_i \Pi_i E\left(vv^T\right) \Pi_i \right) = \sigma_0^2 Tr\left(\Pi_i \left(I \otimes \sigma_0^2 \frac{P_j}{\sigma_j^2} \right) \Pi_i \Pi_i Q_{vv} \Pi_i \right) =$$

$$= \sigma_0^2 \left(Tr\Pi_i - Tr\left(\Pi_i \left(I \otimes \frac{P_j}{\sigma_j^2} \right) \Pi_i \Pi_i A \left(A^T \left(I \otimes \frac{P_j}{\sigma_j^2} \right) A \right)^{-1} A^T \Pi_i \right) \right) = \sigma_0^2 (m_i - n_i)$$

$$\frac{\nu_i\sigma_i^2}{\sigma_0^2} = m_i - n_i = \sum_{k=1}^{m_i} p_k q_{\hat{v}_k} \qquad \text{where:} \qquad q_{\hat{v}_i} = \frac{1}{p_i} - a_i \left(A^T PA \right)^{-1} a_i^T$$

: the local redundancy being: $\nu_i = p_i q_{\hat{v}_i}$ where: $\nu = \sum_{i=1}^{m} \nu_i = m - n$

$$\max_{i=1,h} \left(\left| \hat{\sigma}_i^2 - \tilde{\sigma}_i^2 \right| \right)$$

$$\hat{\sigma}_i^2 = \frac{(\Pi_i\hat{v})^T \Pi_i (I \otimes P_i) \Pi_i (\Pi_i\hat{v})}{(m_i - n_i)} \qquad\qquad i = 1,h$$

- Norm (different from Max Likelihood: L estimators and Least squares: LS suitable for anomalous data and gross error detection by robust estimators)

$$\phi = \left\| p\hat{v} \right\|_k = \left\| qp\hat{v} \right\| = min \qquad \text{auxiliary weights:} \qquad q = \frac{1}{\left(|v|^{2-k} + \varepsilon \right)}$$

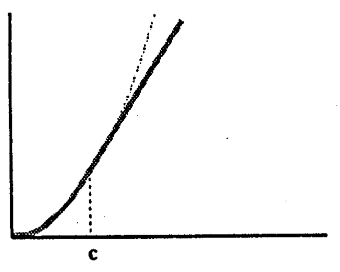
<u>Huber estimator</u> (M estimators: minimax): $\psi(x) = max(-k, min(k, x)) \qquad k > 0$

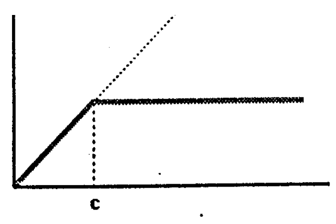
Objective function:

$$\phi(x) = \begin{cases} \frac{1}{2} x^2 & \|x\| < c \\ c\|x\| - \frac{1}{2} c^2 & \|x\| \geq c \end{cases}$$

Influence function:

$$\psi(x) = \begin{cases} x & \|x\| < c \\ c \; sign(x) & \|x\| \geq c \end{cases}$$





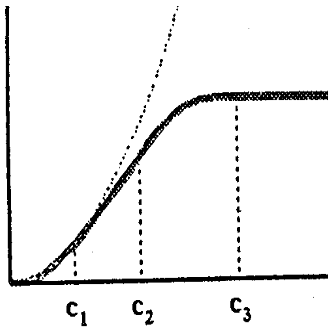
Hampel estimator (S estimators):

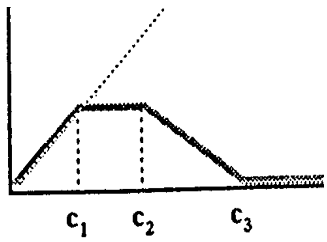
Objective function:

$$\phi(x)=\begin{cases}\frac{1}{2}x^2 & \|x\|<c_1\\ c_1\|x\|-\frac{1}{2}c_1^2 & c_1\le\|x\|\le c_2\\ c_1c_2-\frac{1}{2}c_1^2+\frac{1}{2}c_1(c_3-c_2)\left[1-\left(\frac{c_3-\|x\|}{c_3-c_2}\right)^2\right] & c_2\le\|x\|\le c_3\\ c_1c_2-\frac{1}{2}c_1^2+\frac{1}{2}c_1(c_3-c_2) & \|x\|\ge c_3\end{cases}$$

Influence function:

$$\psi(x)=\begin{cases}x & \|x\|<c_1\\ c_1\ sign(x) & c_1\le\|x\|\le c_2\\ \frac{c_1}{c_3-c_2}(c_3-\|x\|)sign(x) & c_2\le\|x\|\le c_3\\ 0 & \|x\|\ge c_3\end{cases}$$





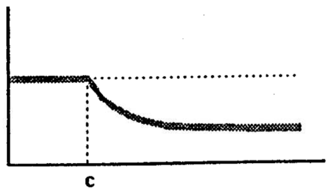
Auxiliary weight function:

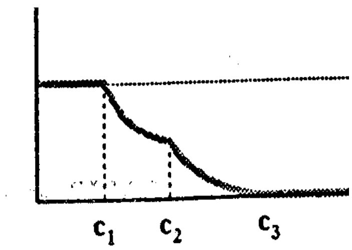
for Huber estimators:

$$q(x)=\begin{cases}1 & \|x\|<c\\ \frac{c}{\|x\|} & \|x\|\ge c\end{cases}$$

for Hampel estimators:

$$q(x)=\begin{cases}1 & \|x\|<c_1\\ \frac{c_1}{\|x\|} & c_1\le\|x\|\le c_2\\ \frac{c_1}{(c_3-c_2)}\left(\frac{c_3}{\|x\|}-1\right) & c_2\le\|x\|\le c_3\\ 0 & \|x\|\ge c_3\end{cases}$$





Numerical tests: **Conditioning (of the parameters)**

- ❑ Singular value decomposition:

$$\varepsilon=\frac{s_{min}}{s_{max}}s_{min}/s_{max}$$

where: $P^{1/2}A=Z=WSY$

- Condition number:

$$\aleph_\infty = \frac{1}{\|C\|_\infty \cdot \|C^{-1}\|_\infty}$$ where: $$\|Q\|_\infty = \max_{i=1,n}\left(\sum_{j=1}^{n}|q_{ij}|\right)$$

$$R = \left(I * C^{-1}\right)^{-1/2} C^{-1}\left(I * C^{-1}\right)^{-1/2}$$

alarm: $\aleph_\infty \le 10^{-3} \div 10^{-5}$ and: $|r_{ij}| \ge 0.7 \div 0.8$

- Conditioning via reliability:

$$\begin{bmatrix} A \\ I \end{bmatrix}[\hat{x}] + \begin{bmatrix} \delta - y_0 \\ -x_0 \end{bmatrix} = \begin{bmatrix} \hat{v} \\ \hat{u} \end{bmatrix} \quad ; \quad \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}$$

$$C_{\hat{u}\hat{u}} = C_{x_0 x_0} - C_{\hat{x}\hat{x}} = \hat{\sigma}_0^2\left(Q^{-1} - Q_{\hat{x}\hat{x}}\right)$$

$$\nu_i = diag\left(Q^{1/2} Q_{\hat{u}\hat{u}} Q^{1/2}\right)_i = diag\left(Q^{1/2}\left(Q^{-1} - Q_{\hat{x}\hat{x}}\right)Q^{1/2}\right)_i = 1 - q_i diag\left(Q_{\hat{x}\hat{x}}\right)_i$$

security: $0 \le 0.20 \div 0.25 \le \nu_i = 1 - q_i diag\left(Q_{\hat{x}\hat{x}}\right) \le 1$

and Reliability (of the observations)

$$\begin{bmatrix} A_1 & 0 \\ A_2 & B \end{bmatrix}\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + \begin{bmatrix} \delta_1 - y_{01} \\ \delta_2 - y_{02} \end{bmatrix} = \begin{bmatrix} \hat{v}_1 \\ \hat{v}_2 \end{bmatrix} \quad ; \quad \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}$$ data set not able to locate outliers

$$\hat{v}_2 \equiv 0 \quad ; \quad \hat{x}_2 = -B^{-1}\left((\delta_2 - y_{02}) + A_2\hat{x}_1\right) \quad ; \quad C_{\hat{v}_2\hat{v}_2} \equiv 0$$

$$\begin{bmatrix} A_1 & 0 \\ A_2 & B \\ A_3 & C \end{bmatrix}\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + \begin{bmatrix} \delta_1 - y_{01} \\ \delta_2 - y_{02} \\ \delta_3 - y_{03} \end{bmatrix} = \begin{bmatrix} \hat{v}_1 \\ \hat{v}_2 \\ \hat{v}_3 \end{bmatrix} \quad ; \quad \begin{bmatrix} P & 0 & 0 \\ 0 & Q & 0 \\ 0 & 0 & R \end{bmatrix}$$ data set able to locate outliers

$$\hat{v}_2' = H\hat{v}_2 \equiv -K\hat{v}_3 = -\hat{v}_3'$$

$$\hat{x}_2 = -\left(B^T QB + C^T RC\right)^{-1}\left(B^T Q\left((\delta_2 - y_{02}) + A_2\hat{x}_1\right) + C^T R\left((\delta_3 - y_{03}) + A_3\hat{x}_1\right)\right)$$
$$\hat{v}_2 = B\left(B^T QB + C^T RC\right)^{-1} C^T RC\left(B^{-1}\left((\delta_2 - y_{02}) + A_2\hat{x}_1\right) - C^{-1}\left((\delta_3 - y_{03}) + A_3\hat{x}_1\right)\right) = -H^{-1}K\hat{v}_3$$
$$\hat{v}_3 = C\left(B^T QB + C^T RC\right)^{-1} B^T QB\left(B^{-1}\left((\delta_2 - y_{02}) + A_2\hat{x}_1\right) - C^{-1}\left((\delta_3 - y_{03}) + A_3\hat{x}_1\right)\right) = -K^{-1}H\hat{v}_2$$

$$C_{\hat{v}_2'\hat{v}_2'} = HC_{\hat{v}_2\hat{v}_2}H^T \equiv KC_{\hat{v}_3\hat{v}_3}K^T = C_{\hat{v}_3'\hat{v}_3'}$$
$$C_{\hat{v}_2\hat{v}_2} = H^{-1}KC_{\hat{v}_3\hat{v}_3}K^T\left(H^{-1}\right)^T \quad ; \quad C_{\hat{v}_3\hat{v}_3} = K^{-1}HC_{\hat{v}_2\hat{v}_2}H^T\left(K^{-1}\right)^T$$

$\hat{v}_2 \equiv -\hat{v}_3$; $C_{\hat{v}_2\hat{v}_2} \equiv C_{\hat{v}_3\hat{v}_3}$ if: $B \equiv C$ and $Q \equiv R$

$$\hat{y} = -\frac{1}{2}B^{-1}\left((\delta_2 - y_{02}) + A_2\hat{x}_1 + (\delta_3 - y_{03}) + A_3\hat{x}_1\right)$$

$$\hat{v}_2 = \frac{1}{2}\left(\left(\left(\delta_2 - y_{02}\right) + A_2\hat{x}_1\right) - \left(\left(\delta_3 - y_{03}\right) + A_3\hat{x}_1\right)\right) \equiv -\hat{v}_3$$

$$\hat{v}_3 = -\frac{1}{2}\left(\left(\left(\delta_2 - y_{02}\right) + A_2\hat{x}_1\right) - \left(\left(\delta_3 - y_{03}\right) + A_3\hat{x}_1\right)\right) \equiv -\hat{v}_2$$

security: $\nu = \dfrac{m+l-n}{n} \geq 2$ and: $0 \leq 0.20 \div 0.25 \leq \nu_i = 1 - p_i a_i Q_{\hat{x}\hat{x}} a_i^T \leq 1$

Internal reliability: $\nabla\left(\delta_i - y_{0i}\right) = \dfrac{\tau\sigma_{y_{0i}}}{\sqrt{\nu_i}}$

External reliability: $\nabla x_j = -\left(\left(A^T PA\right)_j^{-1}\right)^T A^T Pe_i \nabla\left(\delta_i - y_{0i}\right) \qquad \forall\ j$

Appendix A – Robust estimators

- Estimator variance: $\sigma^2(\psi, F) = V(\theta; F) = \dfrac{1}{I(F)}$ where: $I(F)$ Fisher information

- Influence function: $IF(x,\theta;F) = \lim\limits_{t\to 0} \dfrac{\theta((1-t)G + tH) - \theta(G)}{t} =$

  $= \dfrac{\partial}{\partial t}\left[\theta((1-t)G + tH)\right]_{t=0}$

  Note: An estimator is robust if its Influence function remain finite at infinite.

- Rejection point (if exists): $x_0 \Rightarrow IF(x > x_0, \theta, F) \equiv 0$

- Gross–error sensitivity: $\gamma = sup_x \left|IF(x,\theta,F)\right|$

- Local shift sensitivity: $\lambda = \dfrac{\sup\limits_{x\neq y}\left|IF(y,\theta,F) - IF(x,\theta,F)\right|}{|y-x|}$

- Breakdown point: $\varepsilon \Rightarrow F(x-\theta) = (1-\varepsilon)\, G\,(x-\theta) + \varepsilon H(x-\theta)$

  with: $\theta \in G(x-\theta)$

  Note: The higher the value of its Breakdown point (ranging from 0 to 1), the more robust an estimator is.

- Least median of squares (LMedS):

$$y = Ax + \delta \qquad (P) \quad \Rightarrow \quad \begin{bmatrix}\hat{y}_1\\ \hat{y}_2\end{bmatrix} = \begin{bmatrix}y_1^0\\ y_2^0\end{bmatrix} + \begin{bmatrix}0\\ \hat{v}_2\end{bmatrix} = \begin{bmatrix}A_1\\ A_2\end{bmatrix}\hat{x} + \begin{bmatrix}\delta_1\\ \delta_2\end{bmatrix} \qquad \left(\begin{bmatrix}P_1 & 0\\ 0 & P_2\end{bmatrix}\right)$$

where: $dim(y_1) = dim(x) = n$ being: n = number of unknowns =

= number of necessary observations

$dim(y_2) = m - n$ m = number of observations

$m - n$ = number of redundant observations

Note: The LMedS certainly leads to a minimum, but the number of its steps rapidly becomes huge.

$$\binom{m}{n} = \frac{m!}{n!(m-n)!} \qquad ln\binom{m}{n} = ln\, m! - ln\, n! - ln\,(m-n)!$$

n	m	$\binom{m}{n}$
5	15	3.000
7	20	80.000
10	30	30.000.000

with: $k! = \sqrt{2\pi k}\left(\frac{k}{e}\right)^k$

- Random sampling (RANSAC): selection of a certain number of suitable LMedS solutions.

  Note: The RANSAC forms a histogram of solutions, where its modal represents the best one.

Appendix B – Partitioned models

- Kalman filter: (forward solution)

$$\hat{x} = \left(A^T PA + B^T QB\right)^{-1}\left(A^T Py_{0_1} + B^T Qy_{02}\right) \qquad C_{\hat{x}\hat{x}} = \sigma_0^2 Q_{\hat{x}\hat{x}} = \sigma_0^2\left(A^T PA + B^T QB\right)^{-1}$$

$$E\left(\begin{vmatrix} y_1 \\ \\ y_2 \end{vmatrix}\right) = \begin{vmatrix} \overline{y_1} \\ \\ \overline{y_2} \end{vmatrix} = \begin{vmatrix} A \\ \\ B \end{vmatrix} |x|$$

$$D\left(\begin{vmatrix} y_1 \\ \\ y_2 \end{vmatrix}\right) = \begin{vmatrix} C_{y_{0_1} y_{0_1}} & 0 \\ 0 & C_{y_{0_2} y_{0_2}} \end{vmatrix} = \sigma_0^2 \begin{vmatrix} Q_{y_{0_1} y_{0_1}} & 0 \\ 0 & Q_{y_{0_2} y_{0_2}} \end{vmatrix} = \sigma_0^2 \begin{vmatrix} P^{-1} & 0 \\ 0 & Q^{-1} \end{vmatrix}$$

previous solution: $\hat{x}_0 = \left(A^T PA\right)^{-1} A^T Py_{0_1}$ $\qquad C_{\hat{x}_0\hat{x}_0} = \sigma_0^2 Q_{\hat{x}_0\hat{x}_0} = \sigma_0^2\left(A^T PA\right)^{-1}$

(backward solution)

$$\hat{x} = \left(A^T PA - B^T QB\right)^{-1}\left(A^T Py_{01} - B^T Qy_{0_2}\right) \qquad C_{\hat{x}\hat{x}} = \sigma_0^2 Q_{\hat{x}\hat{x}} = \sigma_0^2\left(A^T PA - B^T QB\right)^{-1}$$

$$E(y_1) = E\left(\begin{vmatrix} \bullet \\ \\ y_2 \end{vmatrix}\right) = \overline{y_1} = \begin{vmatrix} \bullet \\ \\ \overline{y_2} \end{vmatrix} = Ax_0 = \begin{vmatrix} \bullet \\ \\ B \end{vmatrix} x_0$$

$$D(y_1) = D\left(\begin{vmatrix} \bullet \\ \\ y_2 \end{vmatrix}\right) = C_{y_{01} y_{01}} = \begin{vmatrix} \bullet & 0 \\ 0 & C_{y_{02} y_{02}} \end{vmatrix} = \sigma_0^2\, Q_{y_{01} y_{01}} = \sigma_0^2 \begin{vmatrix} \bullet & 0 \\ 0 & Q_{y_{0_2} y_{0_2}} \end{vmatrix} = \sigma_0^2\, P^{-1} = \sigma_0^2 \begin{vmatrix} \bullet & 0 \\ 0 & Q^{-1} \end{vmatrix}$$

previous solution: $\hat{x}_0 = \left(A^T PA\right)^{-1} A^T Py_{01}$ $\qquad C_{\hat{x}_0\hat{x}_0} = \sigma_0^2 Q_{\hat{x}_0\hat{x}_0} = \sigma_0^2\left(A^T PA\right)^{-1}$

$$\hat{x} = \left(A^T PA \pm B^T QB\right)^{-1}\left(A^T Py_{0_1} \pm B^T Qy_{02}\right) =$$

$$= \left(Q_{\hat{x}_0\hat{x}_0}^{-1} \pm B^T Q_{y_{02}y_{02}}^{-1} B\right)^{-1}\left(A^T P y_{01} \pm B^T Q y_{02}\right) =$$

$$= \left(Q_{\hat{x}_0\hat{x}_0} \mp Q_{\hat{x}_0\hat{x}_0} B^T \left(Q_{y_{0_2}y_{0_2}} \pm B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1} B Q_{\hat{x}_0\hat{x}_0}\right)\left(A^T P y_{0_1} \pm B^T Q y_{0_2}\right)$$

being: $\mp \cdot \pm = -$

$$\hat{x} = Q_{\hat{x}_0\hat{x}_0} A^T P y_{0_1} \mp Q_{\hat{x}_0\hat{x}_0} B^T \left(Q_{y_{0_2}y_{0_2}} \pm B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1} B Q_{\hat{x}_0\hat{x}_0} A^T P y_{0_1} +$$

$$\pm \left(Q_{\hat{x}_0\hat{x}_0} - Q_{\hat{x}_0\hat{x}_0} B^T \left(Q_{y_{0_2}y_{0_2}} \pm B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1} B Q_{\hat{x}_0\hat{x}_0}\right) B^T Q y_{0_2} =$$

$$= \hat{x}_0 \mp Q_{\hat{x}_0\hat{x}_0} B^T \left(Q_{y_{0_2}y_{0_2}} \pm B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1} B\hat{x}_0 \pm \left(Q_{\hat{x}_0\hat{x}_0} - Q_{\hat{x}_0\hat{x}_0} B^T \left(\left(B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1} +\right.\right.$$

$$\left.\left.\mp \left(B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1}\left(Q_{y_{0_2}y_{0_2}}^{-1} \pm \left(B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1}\right)^{-1}\left(B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1}\right) B Q_{\hat{x}_0\hat{x}_0}\right) B^T Q_{y_{0_2}y_{0_2}}^{-1} y_{0_2}$$

grouping forward: $Q_{\hat{x}_0\hat{x}_0} B^T$ and backward: $B Q_{\hat{x}_0\hat{x}_0} B^T$

being: $A^{-1}B^{-1}C^{-1} = (CBA)^{-1}$ and $A\left(A^{-1} \pm B^{-1}\right)B = B \pm A = \pm(A \pm B)$

moreover: $\pm \cdot - \cdot \mp \cdot \pm = \pm$

$$\hat{x} = \hat{x}_0 \mp Q_{\hat{x}_0\hat{x}_0} B^T \left(Q_{y_{0_2}y_{0_2}} \pm B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1} \hat{y}_2 \pm Q_{\hat{x}_0\hat{x}_0} B^T \left(I +\right.$$

$$\left.-\left(I \mp \left(B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1}\left(Q_{y_{0_2}y_{0_2}}^{-1} \pm \left(B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1}\right)^{-1}\right)\left(B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1} B Q_{\hat{x}_0\hat{x}_0} B^T\right) Q_{y_{0_2}y_{0_2}}^{-1} y_{02} =$$

$$= \hat{x}_0 \mp Q_{\hat{x}_0\hat{x}_0} B^T \left(Q_{y_{0_2}y_{0_2}} \pm B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1} \hat{y}_2 +$$

$$\pm Q_{\hat{x}_0\hat{x}_0} B^T \left(Q_{y_{0_2}y_{0_2}} \left(Q_{y_{0_2}y_{0_2}}^{-1} \pm \left(B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1}\right) B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1} y_{0_2} =$$

$$= \hat{x}_0 \mp Q_{\hat{x}_0\hat{x}_0} B^T \left(Q_{y_{0_2}y_{0_2}} \pm B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1} \hat{y}_2 \pm Q_{\hat{x}_0\hat{x}_0} B^T \left(Q_{y_{0_2}y_{0_2}} \pm B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1} y_{0_2}$$

$$\hat{x} = \hat{x}_0 \pm Q_{\hat{x}_0\hat{x}_0} B^T \left(Q_{y_{0_2}y_{0_2}} \pm Q_{\hat{y}_2\hat{y}_2}\right)^{-1}\left(y_{0_2} - \hat{y}_2\right) = \hat{x}_0 \pm H\hat{v} = \hat{x}_0 \pm \hat{w}$$

where: $w =$ innovation

$$C_{\hat{x}\hat{x}} = \sigma_0^2\left(A^T P A \pm B^T Q B\right)^{-1} = \sigma_0^2\left(Q_{\hat{x}_0\hat{x}_0}^{-1} \pm B^T Q_{y_{0_2}y_{0_2}}^{-1} B\right)^{-1} =$$

$$= \sigma_0^2 Q_{\hat{x}_0\hat{x}_0} \mp \sigma_0^2 Q_{\hat{x}_0\hat{x}_0} B^T \left(Q_{y_{0_2}y_{0_2}} \pm B Q_{\hat{x}_0\hat{x}_0} B^T\right)^{-1} B Q_{\hat{x}_0\hat{x}_0} =$$

$$= \sigma_0^2 Q_{\hat{x}_0\hat{x}_0} \mp \sigma_0^2 Q_{\hat{x}_0\hat{x}_0} B^T \left(Q_{y_{02}y_{02}} \pm Q_{\hat{y}_2\hat{y}_2}\right)^{-1} B Q_{\hat{x}_0\hat{x}_0} = \sigma_0^2 Q_{\hat{x}_0\hat{x}_0} \mp \sigma_0^2 Q_{\hat{x}_0\hat{x}_0} B^T Q_{\hat{v}\hat{v}}^{-1} B Q_{\hat{x}_0\hat{x}_0} =$$

$$= \sigma_0^2 Q_{\hat{x}_0\hat{x}_0} \mp \sigma_0^2 Q_{\hat{x}_0\hat{x}_0} B^T Q_{\hat{v}\hat{v}}^{-1} Q_{\hat{v}\hat{v}} Q_{\hat{v}\hat{v}}^{-1} B Q_{\hat{x}_0\hat{x}_0} = \sigma_0^2 Q_{\hat{x}_0\hat{x}_0} \mp \sigma_0^2 H Q_{\hat{v}\hat{v}} H^T = \sigma_0^2 Q_{\hat{x}_0\hat{x}_0} \mp \sigma_0^2 Q_{\hat{w}\hat{w}}$$

- Gauss blocking: (forward solution)

$$\begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

previous solution: $x_{0_1} = A^{-1} y_1$ (A^{-1})

$$\begin{vmatrix} A & B \\ B^T & C \end{vmatrix} \begin{vmatrix} \alpha & \beta \\ \beta^T & \gamma \end{vmatrix} = \begin{vmatrix} I & 0 \\ 0 & I \end{vmatrix}$$

$$\begin{cases} \alpha = A^{-1} + A^{-1}B\left(C - B^T A^{-1} B\right)^{-1} B^T A^{-1} = \\ \quad = A^{-1} + A^{-1} B \gamma B^T A^{-1} \\ \beta = -A^{-1}B\left(C - B^T A^{-1}B\right)^{-1} = -A^{-1}B\gamma \\ \gamma = \left(C - B^T A^{-1} B\right)^{-1} \end{cases} \qquad \Rightarrow \qquad A^{-1}B = -\beta\gamma^{-1}$$

$$\begin{vmatrix} x_1 \\ x_2 \end{vmatrix} = \begin{vmatrix} \alpha & \beta \\ \beta^T & \gamma \end{vmatrix} \begin{vmatrix} y_1 \\ y_2 \end{vmatrix} = \begin{vmatrix} A^{-1}y_1 + A^{-1}B\left(C - B^T A^{-1}B\right)^{-1}\left(B^T A^{-1} y_1 - y_2\right) \\ -\left(C - B^T A^{-1}B\right)^{-1}\left(B^T A^{-1} y_1 - y_2\right) \end{vmatrix} = \begin{vmatrix} x_{0_1} - A^{-1}By_2 \\ -\left(C - B^T A^{-1}B\right)^{-1}\left(B^T x_{0_1} - y_2\right) \end{vmatrix}$$

$$x = \begin{vmatrix} x_1 \\ x_2 \end{vmatrix}$$

(backward solution)

$$x_{0_1} = x_1 + A^{-1}By_2 = x_1 - \beta\gamma^{-1}y_2 \qquad A^{-1} = \alpha - A^{-1}B\gamma B^T A^{-1} = \alpha - \beta\gamma^{-1}\beta^T$$

Appendix C – Sampling

<u>Note:</u> Bernoullian sampling is strictly required.

- Logistic strategies:

(logistic function)

$$y = a + (b-a)e^{\frac{-4(b-a)}{e^2 cx}} \qquad lim_{x\to 0+}\, y = a \qquad lim_{x\to\infty}\, y = b$$

$$y^I = (b-a)e^{\frac{-4(b-a)}{e^2 cx}}\,\frac{4(b-a)}{e^2 cx^2}$$

$$y^{II} = (b-a)c^{\frac{-4(b-a)}{e^2 cx}}\,\frac{b(b-a)^2}{e^4 c^2 x^4}\left(1 - \frac{e^2 cx}{2(b-a)}\right)$$

$$y^{II} = 0 \qquad x = \frac{2(b-a)}{e^2 c} = K$$

$$y(K) = a + (b-a)e^{-2} \qquad y^I(K) = c$$

<u>for the test of frequency:</u>

$$z_{inf} = \frac{x_{inf} + 0.5 - n\hat{p}}{\sqrt{n\hat{p}q}} \qquad z_{sup} = \frac{x_{sup} - 0.5 - n\hat{p}}{\sqrt{n\hat{p}q}}$$

$$z_{inf}\sqrt{\frac{\hat{p}q}{n}} = \frac{x_{inf}}{n} + \frac{1}{2n} - \hat{p} = \pi f_{inf} + \frac{1}{2n} - \hat{p}$$

$$z_{sup}\sqrt{\frac{\hat{p}q}{n}} = \frac{x_{sup}}{n} - \frac{1}{2n} - \hat{p} = \pi f_{sup} - \frac{1}{2n} - \hat{p}$$

$$f_{inf} = \left(\hat{p} - \frac{1}{2n}\right) - z_\alpha\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \qquad f_{sup} = \left(\hat{p} + \frac{1}{2n}\right) + z_\alpha\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

where: $z_\alpha = z_{sup} = -z_{inf}$

$$f_{inf} = \left(\hat{p} - \frac{1}{2n}\right) - z_\alpha \sqrt{\frac{\hat{p}(1-\hat{p})}{n}\frac{N-n}{N-1}} \qquad f_{sup} = \left(\hat{p} + \frac{1}{2n}\right) + z_\alpha \sqrt{\frac{\hat{p}(1-\hat{p})}{n}\frac{N-n}{N-1}}$$

$$f_{sup} - f_{inf} = \frac{1}{n} + 2\, z_\alpha \sqrt{\frac{\hat{p}(1-\hat{p})}{n}\frac{N-n}{N-1}} = 2a$$

$$2\, z_\alpha \sqrt{\frac{\hat{p}(1-\hat{p})}{n}\frac{N-n}{N}} = 2a - \frac{1}{n} \qquad 4z_\alpha^2 n\left[\hat{p}(1-\hat{p})\frac{N-n}{N}\right] = 4a^2n^2 + 1 - 4an$$

$$4\left[a^2 + z_\alpha^2 \hat{p}\frac{(1-\hat{p})}{N}\right]n^2 - 4\left[a + z_\alpha^2 \hat{p}(1-\hat{p})\right]n + 1 = 0$$

$$n = \frac{\left[a + z_\alpha^2 \hat{p}(1-\hat{p})\right] \pm \sqrt{z_\alpha^4 \hat{p}^2(1-\hat{p})^2 + z_{\alpha\alpha}^2 \hat{p}(1-\hat{p})\left(2a - \frac{1}{N}\right)}}{2\left(a + z_\alpha^2 \hat{p}\frac{(1-\hat{p})}{N}\right)}$$

$$n = \frac{(a+1) \pm \sqrt{1 + 1\left(2a - \frac{1}{N}\right)}}{2\left(a + \frac{1}{N}\right)} = \frac{N(a+1) \pm \sqrt{1+2a}}{2\left(a + \frac{1}{N}\right)} = \frac{N\left[(a+1) \pm (a+1)\right]}{2\left(1 + Na^2\right)}$$

$$n = \frac{N(a+1)}{1 + Na^2} = \frac{N}{1 + Na^2}$$

<u>for the test of variance:</u>

$$\hat{\sigma}^2 + \chi_{sup}^2 \frac{\hat{\sigma}^2}{\sqrt{2n}} - \left(\hat{\sigma}^2 - \chi_{inf}^2 \frac{\hat{\sigma}^2}{\sqrt{2n}}\right) = 2a \qquad \left(\chi_{sup}^2 + \chi_{inf}^2\right)\frac{\hat{\sigma}^2}{\sqrt{2n}} = 2a$$

where: $\hat{\sigma}_{\hat{\sigma}^2}^2 = \dfrac{\hat{\sigma}^4}{2N}$

$$n = \frac{\left(\chi_{sup}^2 + \chi_{inf}^2\right)^2 \hat{\sigma}^4}{8a^2}$$

$$\hat{\sigma}^2 + \chi_{sup}^2 \frac{\hat{\sigma}^2}{\sqrt{2n}}\frac{\sqrt{N-n}}{\sqrt{N}} - \left(\hat{\sigma}^2 - \chi_{inf}^2 \frac{\hat{\sigma}^2}{\sqrt{2n}}\frac{\sqrt{N-n}}{\sqrt{N}}\right) = 2a \qquad \left(\chi_{sup}^2 + \chi_{inf}^2\right)\frac{\hat{\sigma}^2}{\sqrt{2n}}\frac{\sqrt{N-n}}{\sqrt{N}} = 2a$$

where: $\hat{\sigma}_{\hat{\sigma}^2}^2 = \dfrac{\hat{\sigma}^4}{2n}\dfrac{N-n}{N-1}$

$$2a\sqrt{N}\sqrt{2n} = \hat{\sigma}^2\left(\chi_{sup}^2 + \chi_{inf}^2\right)\sqrt{N-n}$$

$$n = \frac{\hat{\sigma}^4 \left(\chi^2_{sup} + \chi^2_{inf}\right)^2 N}{8a^2 N + \hat{\sigma}^4 \left(\chi^2_{sup} + \chi^2_{inf}\right)^2} \qquad n = \frac{\nu^2 N}{\nu^2 + 2a^2 N}$$

where: $E(\chi^2) = \nu$ and: $\chi^2_{sup} + \chi^2_{inf} \cong 2\nu$

- Stratified sampling:

$$\sigma^2_\mu = \frac{\bar{\sigma}^2_x}{n} = \frac{(n-1)\sigma^2_x}{n(n-m)} - \frac{(m-1)\sigma^2_{\bar{x}}}{n(n-m)} < \frac{\sigma^2_x}{n}$$

$$\sigma^2_\mu = \frac{\bar{\sigma}^2_x}{n}\frac{N-n}{N-1} = \left(\frac{(n-1)\sigma^2_x}{n(n-m)} - \frac{(m-1)\sigma^2_{\bar{x}}}{n(n-m)}\right)\frac{N-n}{N-1} < \frac{\sigma^2_x}{n}\frac{N-n}{N-1}$$

where: $\sigma^2_x = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ $\qquad \sigma^2_{\bar{x}} = \frac{m}{m-1}\sum_{i=1}^{m} p_i (\bar{x}_i - \bar{x})^2$

$$\bar{\sigma}^2_x = \frac{(n-1)\sigma^2_x - (m-1)\sigma^2_{\bar{x}}}{n-m}$$

$$\sigma^2_\mu = \frac{\sigma^2_\sigma}{n} = \frac{\bar{\sigma}^2_x}{n} - \frac{\sigma^2_0}{n} < \frac{\bar{\sigma}^2_x}{n} < \frac{\sigma^2_x}{n}$$

$$\sigma^2_\mu = \frac{\sigma^2_\sigma}{n}\frac{N-n}{N-1} = \left(\frac{\bar{\sigma}^2_x}{n} - \frac{\sigma^2_0}{n}\right)\frac{N-n}{N-1} < \frac{\bar{\sigma}^2_x}{n}\frac{N-n}{N-1} < \frac{\sigma^2_x}{n}\frac{N-n}{N-1}$$

where: $p_i = \sigma_i / \sigma_0$ $\qquad \sigma_0 = \sum p_i \sigma_i$ $\qquad \bar{\sigma}^2_x = \sum p_i \sigma^2_i$

$$n_i = nN_i\sigma_i / (N\sigma_0) = np_i \qquad \sigma^2_\sigma = \sum_{i=1}^{m} p_i (\sigma_i - \sigma_0)^2 = \sum_{i=1}^{m} p_i \sigma^2_i - \sigma^2_0 = \bar{\sigma}^2_x - \sigma^2_0$$

Appendix D – Design of experiments

Optimization order:	Parameters: known	Parameters: unknown
Zero	$A,\ P$	Q_{xx}
First	$P,\ Q_{xx}$	A
Second	$A,\ Q_{xx}$	P
Third	Q_{xx}	A, P (partially)

- Methods:
  - simulation;
  - sensitivity analysis;
  - optimal design;
  - improved design;
  - Monte Carlo.
- Numerical tests:
  - conditioning;
  - accuracy;
  - precision;
  - reliability;
  - robustness.

PART III – MULTIVARIATE ANALYSIS

CLUSTER ANALYSIS

Elements and characteristics:

clusters (output data)
cluster elements (input data)
clustering or clumping [6] strategies
agglomerative or divisive or sequential [7] techniques
cluster points (centroids)

Conditions:

$$C_k \cap C_h = 0 \qquad \forall k,h$$

$$C_k \neq 0 \qquad \forall k$$

$$C_1 \cup C_2 \cup ... \cup C_{n-1} \cup C_n = S$$

Target functions:

$$\phi_{p,q,r} = \sum_{k=1}^{n} \sum_{i=1}^{m} u_{ik}{}^{r} \left\| \underline{x}_i - \underline{y}_k \right\|_p^q \qquad 1 \le p \le \infty \; ; \; q \ge 1 \; ; \; r \ge 1$$

$$\phi_{p,q,r} = \sum_{k=1}^{n(j)} \sum_{i=1}^{m} u_{ik}^{r} \left\| \underline{x}_i - \underline{y}_k \right\|_p^q \qquad 1 \le p \le \infty \; ; \; q \ge 1 \; ; \; r \ge 1$$ [8]

$$\sum_{i=1}^{m} u_{ik}^{r} \left\| \underline{x}_i - \underline{y}_k \right\|_p^q \le s_{max} \qquad \forall k \; ; \; k = 1,n(j)$$

$$\left\| \underline{y}_k - \underline{y}_l \right\|_p^q \ge d_{min} \qquad \forall k,l \; ; \; k,l = 1,n(j)$$

$$\phi_{p,q,r} = \frac{\sum_{k=1}^{n} \sum_{i=1}^{m} \sum_{j=1}^{m} u_{ik}^{r} u_{jk}^{r} \left\| \underline{x}_i - \underline{x}_j \right\|_p^q}{2 \sum_{i=1}^{m} u_{ik}^{r}} \qquad 1 \le p \le \infty \; ; \; q \ge 1 \; ; \; r \ge 1$$

$$\phi_{q,r} = \sum_{k=1}^{n} \sum_{i=1}^{m} u_{ik}^{r} D^{q}(i,k) \qquad q \ge 1 \; ; \; r \ge 1$$ [9]

$$\phi_{q,r} = \frac{\sum_{k=1}^{n} \sum_{i=1}^{m} \sum_{j=1}^{m} u_{ik}^{r} u_{jk}^{r} d^{q}(i,j)}{2 \sum_{i=1}^{m} u_{ik}^{r}} \qquad q \ge 1 \; ; \; r \ge 1$$

[6] Clumping Techniques introduce to Fuzzy Sets and Membership Functions.
[7] Sequential Techniques are called: split and merge.
[8] Accepting a free number of clusters, for numerical data only.
[9] Using "dissimilarities" for non-numerical data.

$0 \le u_{ik} \le 1 \qquad \forall i,k$ (fuzzy methods)

$u_{ik} = 0 \;\; ; \;\; 1 \qquad \forall i,k$ (binary methods)

$$\sum_{k=1}^{n} u_{ik} = 1 \qquad \forall i$$

MULTIPLE REGRESSION

Input:

Parameters / *Observations*	x_1	x_2	...	x_i	...	x_{i*}	...	x_n
1	x_{11}	x_{12}	...	x_{1i}	...	x_{1i*}	...	x_{1n}
2	x_{21}	x_{22}	...	x_{2i}	...	x_{2i*}	...	x_{2n}
...	...	...	...	...	...	...	...	...
k	x_{k1}	x_{k2}	...	x_{ki}	...	x_{ki*}	...	x_{kn}
...	...	...	...	...	...	...	...	...
m	x_{m1}	x_{m2}	...	x_{mi}	...	x_{mi*}	...	x_{mn}

Mean values $\quad \bar{x}_1 \quad \bar{x}_2 \quad \dots \quad \bar{x}_i \quad \dots \quad \bar{x}_{i*} \quad \dots \quad \bar{x}_n$

Variances $\quad \sigma^2_{x_1} \quad \sigma^2_{x_2} \quad \dots \quad \sigma^2_{x_i} \quad \dots \quad \sigma^2_{x_{i*}} \quad \dots \quad \sigma^2_{x_n}$

$$i* \quad \Rightarrow \quad x_{ki*} = y_k \qquad (k = 1,m)$$

System:

$$y_k = b_0 + \sum_{(i \ne i*)} {}_i\, b_i x_{ki} \qquad \forall k$$

$$v_k = y_k - y_k^0 = b_0 + \sum_{(i \ne i*)} {}_i\, b_i x_{ki} - y_k^0 \qquad \forall k$$

$$w_k = \sum_{(i \ne i*)} {}_i\, a_i x''_{ki} - y''_k \qquad \forall k$$

being: $\quad x'_{ki} = x_{ki} - \bar{x}_i \qquad x''_{ki} = x'_{ki} \big/ \sigma_{x\ i}$

$y'_k = y_k - \bar{y} \qquad y''_k = y'_k \big/ \sigma_y \qquad a_i = b_i \sigma_{xi} \big/ \sigma_y$

$$\sum_k w_k^2 = min \qquad \sum_k w_k x''_{ki} = 0$$

$$\sum_{(i \ne i*)} {}_i\, a_i \sum_k x''_{ki}\, x''_{kj} = \sum_k y''_k x''_{kj} \qquad \forall j \ne i^*$$

Strategies:

	x_1	x_2	$\cdots$	x_j	$\cdots$	x_{n-1}	y
x_1	r_{11}	r_{12}	$\cdots$	r_{1j}	$\cdots$	$r_{1(n-1)}$	r_{1y}
x_2	r_{21}	r_{22}	$\cdots$	r_{2j}	$\cdots$	$r_{2(n-1)}$	r_{2y}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_i	r_{i1}	r_{i2}	$\cdots$	r_{ii}	$\cdots$	$r_{i(n-1)}$	r_{iy}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_{n-1}	$r_{(n-1)1}$	$r_{(n-1)2}$	$\cdots$	$r_{(n-1)j}$	$\cdots$	$r_{(n-1)(n-1)}$	$r_{(n-1)y}$
y	r_{y1}	r_{y2}	$\cdots$	r_{yj}	$\cdots$	$r_{y(n-1)}$	r_{yy}

Orthogonal variance decomposition and error minimization:

$$S_T^2 = \sum_k (y_k^0 - \bar{y})^2 = \sum_k \left(y_k^0 - y_k\right)^2 + \sum_k \left(y_k - \bar{y}\right)^2 + 2\sum_k \left(y_k^0 - y_k\right)\left(y_k - \bar{y}\right) =$$

$$= S_R^2 + S_S^2 - 2\sum_k v_k \sum_i b_i x'_{ki} = S_R^2 + S_S^2 - 2\sum_i b_i \sum_k x'_{ki} v_k = S_R^2 + S_S^2 \qquad \underline{\text{Note:}} \quad \sum_i b_i \sum_k x'_{ki} v_k \equiv 0$$

$$S_R^2 = \sum_k \left(y_k^0 - y_k\right)^2 = \left(m-1-n'\right)\sigma_y^2 r_{yy}^{(l)} = \left(m-1-n'\right)\sigma_y^2 \, S_R^{2(l)} \big/ S_T^2 = min$$

$$\Rightarrow \qquad \Delta S_S^2 \big/ S_T^2 = r_{yh}^{(l)} r_{hy}^{(l)} \big/ r_{hh}^{(l)} = max$$

Computational techniques:

$element\ (h,h)$	$r_{hh}^{(l+1)} = 1 \big/ r_{hh}^{(l)}$	
$row\ (h,.)$	$r_{hj}^{(l+1)} = r_{hj}^{(l)} \big/ r_{hh}^{(l)}$	$r_{hy}^{(l+1)} = r_{hy}^{(l)} \big/ r_{hh}^{(l)}$
$colunm\ (.,h)$	$r_{ih}^{(l+1)} = -\, r_{ih}^{(l)} \big/ r_{hh}^{(l)}$	$r_{yh}^{(l+1)} = -\, r_{yh}^{(l)} \big/ r_{hh}^{(l)}$
$other\ elements\ (.,.)$	$r_{ij}^{(l+1)} = r_{ij}^{(l)} - r_{ih}^{(l)} r_{hj}^{(l)} \big/ r_{hh}^{(l)}$	$r_{iy}^{(l+1)} = r_{iy}^{(l)} - r_{ih}^{(l)} r_{hy}^{(l)} \big/ r_{hh}^{(l)}$
		$r_{yj}^{(l+1)} = r_{yj}^{(l)} - r_{yh}^{(l)} r_{hj}^{(l)} \big/ r_{hh}^{(l)}$
$taget\ \ function$		$r_{yy}^{(l+1)} = r_{yy}^{(l)} - r_{yh}^{(l)} r_{hy}^{(l)} \big/ r_{hh}^{(l)}$

Fisher tests:

Outgoing variable: $$F^{(l+1)} = \upsilon_2 \left(-\, r_{yh}^{(l)} r_{hy}^{(l)} \big/ r_{hh}^{(l)} \right) \big/ r_{yy}^{(l)} =$$

$$= \upsilon_2 \left(\Delta S_S^2 \big/ S_T^2 \right) \big/ \left(S_R^2 \big/ S_T^2 \right) = \upsilon_2 \Delta S_S^2 \,/\, S_R^2$$

Incoming variable: $F^{(l+1)} = (\upsilon_2 - 1)(r_{yh}^{(l)} r_{hy}^{(l)} / r_{hh}^{(l)}) \big/ (r_{yy}^{(l)} - r_{yh}^{(l)} r_{hy}^{(l)} \big/ r_{hh}^{(l)}) =$

$$= (\upsilon_2 - 1)(\Delta S_S^2 \big/ S_T^2) \big/ (S_R^2 \big/ S_T^2) = (\upsilon_2 - 1)\Delta S_S^2 \big/ S_R^2$$

Output:

$$y_k = \bar{y} + \sum_i b_i (x_{ki} - \bar{x}_i) \qquad \sigma_{y_k}^2 = \sigma_{\bar{y}}^2 + \sum_i \sigma_{b_i}^2 (x_{ki} - \bar{x}_i)^2 + 2\sum_{\substack{ij \\ j>i}} \sigma_{b_i b_j} (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j) \qquad \forall k$$

$$v_k = y_k - y_k^0 = b_0 + \sum_{(i \neq i^*)} {}_i b_i x_{ki} - y_k^0 \qquad \forall k \qquad\qquad \sigma_0 = \sqrt{\frac{S_R^2}{m-1-n'}} = \sqrt{\frac{\sum_k v_k^2}{m-1-n'}}$$

$y_k^{+/-} = y_k \pm t_{\alpha/2} \sigma_{y_k}$ (confidence bound)

Regression validation: $R^{(1)} = S_S^2 \big/ S_T^2 = 1 - S_R^2 \big/ S_T^2 = 1 - r_{yy}^{(1)}$

$$R_{yh}^{(1)} = -\Delta S_S^{2(1)} \big/ S_R^{2\,(1)} = (-r_{yh}^{(1)} r_{hy}^{(1)} \big/ r_{hh}^{(1)}) \big/ r_{yy}^{(1)} \qquad \forall h$$

ANALYSIS OF VARIANCE (ANOVA)

One way:

$$\bar{a} + \hat{a}_i = s_{ij}^o + \hat{v}_{ij} \qquad \forall i = 1,....,I \qquad \forall j = 1,....,J$$

$\hat{a}_I \equiv 0$ (constraint)

Two ways:

$$\bar{a} + \hat{a}_i + \hat{a}_j = s_{ijk}^o + \hat{v}_{ijk} \qquad \forall i = 1,....,I \qquad \forall j = 1,....,J \qquad \forall k = 1,....,K$$

$\hat{a}_I \equiv 0, \quad \hat{a}_J \equiv 0$ (constraints)

Two ways with interaction:

$$\bar{a} + \hat{a}_i + \hat{a}_j + \hat{a}_{ij} = s_{ijk}^o + \hat{v}_{ijk} \qquad \forall i = 1,....,I \qquad \forall j = 1,....,J \qquad \forall k = 1,....,K$$

$\hat{a}_I \equiv 0, \quad \hat{a}_J \equiv 0 \quad \hat{a}_{iJ} \equiv 0 \quad \forall i, \quad \hat{a}_{Ij} \equiv 0 \quad \forall j \neq J$ (constraints)

Fisher tests [10]

$$\sigma_T^2 = \frac{1}{I-1} \sum_{i=1}^{I} \hat{a}_i^2 \qquad\qquad \sigma_B^2 = \frac{1}{J-1} \sum_{j=1}^{J} \hat{a}_j^2$$

$$\sigma_{TB}^2 = \frac{1}{(I-1)(J-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} \hat{a}_{ij}^2$$

[10] The test of sign represents an alternative and it permits to compare not-independent estimates.

$$\sigma_R^2 = \frac{1}{\nu}\sum_{i=1}^{I}\sum_{j=1}^{J}\hat{v}_{ij}^2 \qquad \text{(one way)} \qquad\qquad \sigma_R^2 = \frac{1}{\nu}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\hat{v}_{ijk}^2 \quad \text{(two ways)}$$

$$\begin{aligned} \nu &= I\cdot J-(I-1)-1 \qquad \text{(degrees of freedom)}\\ \nu &= I\cdot J\cdot K-(I-1)-(J-1)-1\\ \nu &= I\cdot J\cdot K-(I-1)-(J-1)-(I-1)(J-1)-1 \end{aligned}$$

Three ways:

$$\overline{a}+\hat{a}_i+\hat{a}_j+\hat{a}_k = s_{ijkl}^o+\hat{v}_{ijkl} \qquad \begin{aligned} &\forall i=1,....,I\\ &\forall j=1,....,J\\ &\forall k=1,....,K\\ &\forall l=1,....,L \end{aligned}$$

$$\hat{a}_I\equiv 0,\qquad \hat{a}_J\equiv 0,\qquad \hat{a}_K\equiv 0 \qquad \text{(constraints)}$$

Three ways with one interaction:

$$\overline{a}+\hat{a}_i+\hat{a}_j+\hat{a}_k+\hat{a}_{ij} = s_{ijkl}^o+\hat{v}_{ijkl} \qquad \begin{aligned} &\forall i=1,....,I\\ &\forall j=1,....,J\\ &\forall k=1,....,K\\ &\forall l=1,....,L \end{aligned}$$

$$\begin{aligned} &\hat{a}_I\equiv 0,\qquad \hat{a}_J\equiv 0,\qquad \hat{a}_K\equiv 0 \qquad \text{(constraints)}\\ &\hat{a}_{iJ}\equiv 0 \qquad \forall i,\qquad \hat{a}_{Ij}\equiv 0 \qquad \forall j\neq J \end{aligned}$$

Three ways with three interactions:

$$\overline{a}+\hat{a}_i+\hat{a}_j+\hat{a}_k+\hat{a}_{ij}+\hat{a}_{ik}+\hat{a}_{jk} = s_{ijkl}^o+\hat{v}_{ijkl} \qquad \begin{aligned} &\forall i=1,....,I\\ &\forall j=1,....,J\\ &\forall k=1,....,K\\ &\forall l=1,....,L \end{aligned}$$

$$\begin{aligned} &\hat{a}_I\equiv 0,\qquad \hat{a}_J\equiv 0,\qquad \hat{a}_K\equiv 0 \qquad \text{(constraints)}\\ &\hat{a}_{iJ}\equiv 0 \qquad \forall i,\qquad \hat{a}_{Ij}\equiv 0 \qquad \forall j\neq J\\ &\hat{a}_{iK}\equiv 0 \qquad \forall i,\qquad \hat{a}_{Ik}\equiv 0 \qquad \forall k\neq K\\ &\hat{a}_{jK}\equiv 0 \qquad \forall j,\qquad \hat{a}_{Jk}\equiv 0 \qquad \forall k\neq K \end{aligned}$$

Three ways with three interactions and a multiple interaction:

$$\overline{a}+\hat{a}_i+\hat{a}_j+\hat{a}_k+\hat{a}_{ij}+\hat{a}_{ik}+\hat{a}_{jk}+\hat{a}_{ijk} = s_{ijkl}^o+\hat{v}_{ijkl} \qquad \begin{aligned} &\forall i=1,....,I\\ &\forall j=1,....,J\\ &\forall k=1,....,K\\ &\forall l=1,....,L \end{aligned}$$

$$\hat{a}_I\equiv 0,\qquad \hat{a}_J\equiv 0,\qquad \hat{a}_K\equiv 0 \qquad \text{(constraints)}$$

$$\hat{a}_{iJ} \equiv 0 \qquad \forall i, \qquad \hat{a}_{Ij} \equiv 0 \qquad \forall j \neq J$$
$$\hat{a}_{iK} \equiv 0 \qquad \forall i, \qquad \hat{a}_{Ik} \equiv 0 \qquad \forall k \neq K$$
$$\hat{a}_{jK} \equiv 0 \qquad \forall j, \qquad \hat{a}_{Jk} \equiv 0 \qquad \forall k \neq K$$
$$\hat{a}_{ijK} \equiv 0 \qquad \forall i \neq I, \qquad \forall j \neq J$$
$$\hat{a}_{iJk} \equiv 0 \qquad \forall i \neq I, \qquad \forall k \neq K$$
$$\hat{a}_{IjK} \equiv 0 \qquad \forall j \neq J, \qquad \forall k \neq K$$
$$\hat{a}_{iJK} \equiv 0 \qquad \forall i \neq I$$
$$\hat{a}_{IjK} \equiv 0 \qquad \forall j \neq J$$
$$\hat{a}_{IJk} \equiv 0 \qquad \forall k \neq K \qquad \hat{a}_{IJK} \equiv 0$$

Fisher tests:

$$\sigma_T^2 = \frac{1}{I-1} \sum_{i=1}^{I} \hat{a}_i^2 \qquad \sigma_B^2 = \frac{1}{J-1} \sum_{j=1}^{J} \hat{a}_j^2 \qquad \sigma_S^2 = \frac{1}{K-1} \sum_{k=1}^{K} \hat{a}_k^2$$

$$\sigma_{TB}^2 = \frac{1}{(I-1)(J-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} \hat{a}_{ij}^2 \qquad \sigma_{TS}^2 = \frac{1}{(I-1)(K-1)} \sum_{i=1}^{I} \sum_{k=1}^{K} \hat{a}_{ik}^2$$
$$\sigma_{BS}^2 = \frac{1}{(J-1)(K-1)} \sum_{j=1}^{J} \sum_{k=1}^{K} \hat{a}_{jk}^2 \qquad \sigma_{TBS}^2 = \frac{1}{(I-1)(J-1)(K-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \hat{a}_{ijk}^2$$

$$\sigma_R^2 = \frac{1}{\nu} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{l=1}^{L} \hat{v}_{ijkl}^2$$

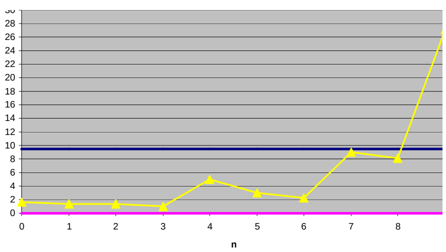
$$\nu = I \cdot J \cdot K \cdot L - (I-1) - (J-1) - (K-1) - 1$$ (degrees of freedom)
$$\nu = I \cdot J \cdot K \cdot L - (I-1) - (J-1) - (K-1) - (I-1)\cdot(J-1) - 1$$
$$\nu = I \cdot J \cdot K \cdot L - (I-1) - (J-1) - (K-1) - (I-1)\cdot(J-1) - (I-1)\cdot(K-1) - (J-1)\cdot(K-1) - 1$$
$$\nu = I \cdot J \cdot K \cdot L - (I-1) - (J-1) - (K-1) - (I-1)\cdot(J-1) - (I-1)\cdot(K-1) - (J-1)\cdot(K-1) + \\ - (I-1)(J-1)(K-1) - 1$$

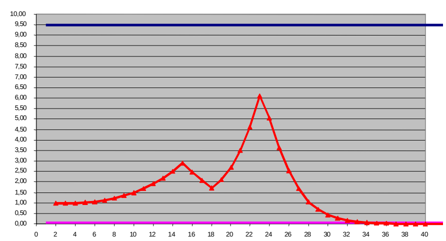
SEQUENTIAL TESTS

Confidence bounds: $\lambda_0 = (1-\alpha)/\beta$ $\qquad$ $\lambda_1 = \alpha/(1-\beta)$

Likelihood ratio: $\lambda = \dfrac{P(x \Rightarrow if\ H_0\ is\ true)}{P(x \Rightarrow if\ H_1\ is\ true)}$ $\qquad$ $\lambda = \prod_{i=1}^{n} p_0(x_i) \Big/ \prod_{i=1}^{n} p_1(x_i)$

$$ln\ \lambda = \sum_{i=1}^{n} ln(p_0(x_i)) - \sum_{i=1}^{n} ln(p_1(x_i))$$





PART IV – INTERPOLATION AND APPROXIMATION

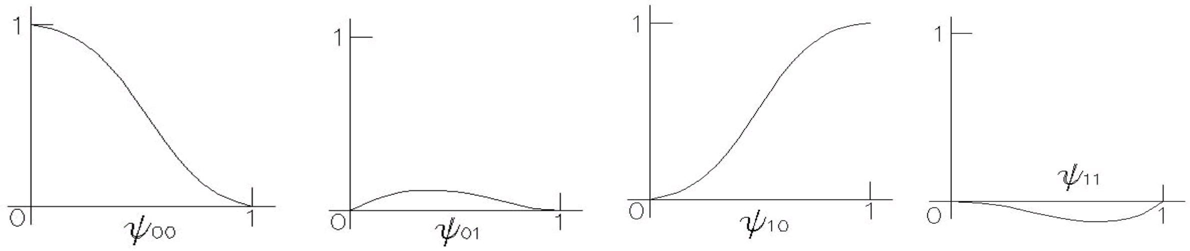
FINITE METHOD INTERPOLATION

$y = 2x^3 - 3x^2 + 1$ (I) *one-dimension cubic case*

$y = x^3 - 2x^2 + x$ (II)

$y = -2x^3 + 3x^2$ (III)

$y = x^3 - x^2$ (IV)



Base functions

two-dimension cubic case

$z = (2x^3 - 3x^2 + 1)(2y^3 - 3y^2 + 1)$ (I.1)

$z = (2x^3 - 3x^2 + 1)(y^3 - 3y^2 + 1)$ (I.2)

$z = (2x^3 - 3x^2 + 1)(-2y^3 + 3y^2)$ (I.3)

$z = (2x^3 - 3x^2 + 1)(y^3 - y^2)$ (I.4)

$z = (x^3 - 3x^2 + x)(2y^3 - 3y^2 + 1)$ (II.1)

$z = (x^3 - 3x^2 + x)(y^3 - 3y^2 + y)$ (II.2)

$z = (x^3 - 3x^2 + x)(-2y^3 + 3y^2)$ (II.3)

$z = (x^3 - 3x^2 + x)(y^3 - y^2)$ (II.4)

$z = (-2x^3 + 3x^2)(2y^3 - 3y^2 + 1)$ (III.1)

$z = (-2x^3 + 3x^2)(y^3 - 3y^2 + y)$ (III.2)

$z = (-2x^3 + 3x^2)(-2y^3 + 3y^2)$ (III.3)

$z = (-2x^3 + 3x^2)(y^3 - y^2)$ (III.4)

$z = (x^3 - x^2)(2y^3 - 3y^2 + 1)$ (IV.1)

$z = (x^3 - x^2)(y^3 - 3y^2 + y)$ (IV.2)

$z = (x^3 - x^2)(-2y^3 + 3y^2)$ (IV.3)

$z = (x^3 - x^2)(y^3 - y^2)$ (IV.4)

SPLINE INTERPOLATION

$f(s_1) = 0$ $\quad s_1 < -2a$ *one-dimension linear case*

$f(s_1) = \dfrac{2a + s_1}{4a^2}$ $\quad -2a \le s_1 \le 0$

$f(s_1) = \dfrac{2a - s_1}{4a^2}$ $\quad 0 \le s_1 \le 2a$

$f(s_1) = 0$ $\quad s_1 > 2a$

$f(s_3) = 0$ $\quad s_3 < -4a$ *one-dimension cubic case*

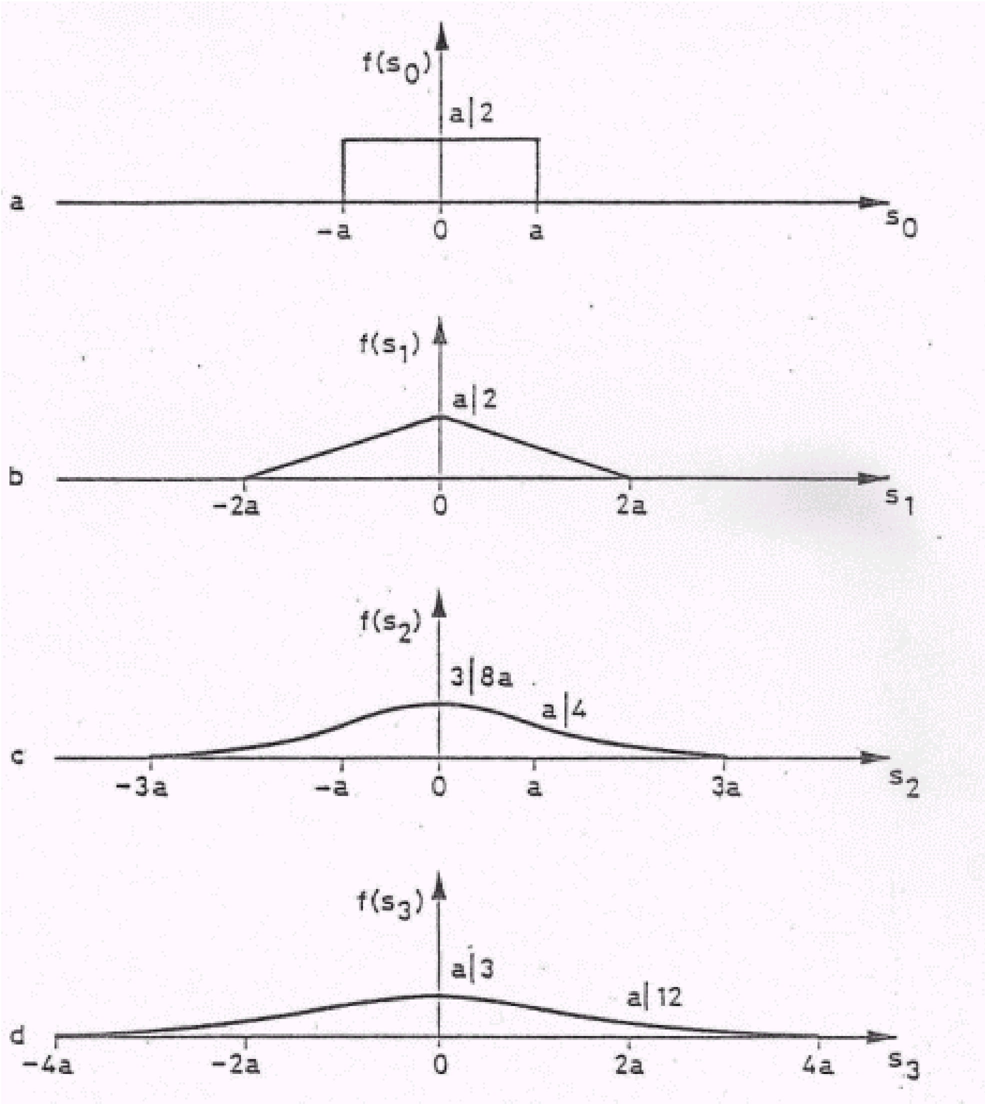
$f(s_3) = \dfrac{(4a + s_3)^3}{96a^4}$ $\quad -4a \le s_3 \le -2a$

$f(s_3) = \dfrac{(4a + s_3)^3 - 4(2a + s_3)^3}{96a^4}$ $\quad -2a \le s_3 \le 0$

$f(s_3) = \dfrac{(4a - s_3)^3 - 4(2a - s_3)^3}{96a^4}$ $\quad 0 \le s_3 \le 2a$

$$f(s_3) = \frac{(4a - s_3)^3}{96a^4} \qquad 2a \le s_3 \le 4a$$

$$f(s_3) = 0 \qquad s_3 > 4a$$



Spline functions of 0 1st 2nd and 3rd order

two-dimension linear case

$$f(s_1,t_1) = \frac{1}{16a^4}(2a + s_1)(2a + t_1) \qquad -2a \le s_1 \le 0 \quad \text{and} \quad -2a \le t_1 \le 0$$

$$f(s_1,t_1) = \frac{1}{16a^4}(2a + s_1)(2a - t_1) \qquad -2a \le s_1 \le 0 \quad \text{and} \quad 0 \le t_1 \le 2a$$

$$f(s_1,t_1) = \frac{1}{16a^4}(2a - s_1)(2a + t_1) \qquad 0 \le s_1 \le 2a \quad \text{and} \quad -2a \le t_1 \le 0$$

$$f(s_1,t_1) = \frac{1}{16a^4}(2a - s_1)(2a - t_1) \qquad 0 \le s_1 \le 2a \quad \text{and} \quad 0 \le t_1 \le 2a$$

two-dimension cubic case

$$f(s_3,t_3) = \frac{1}{9216a^8}(4a + s_3)^3(4a + t_3)^3 \qquad -4a \le s_3 \le -2a \quad \text{and} \quad -4a \le t_3 \le -2a$$

$$f(s_3,t_3)=\frac{1}{9216a^8}(4a+s_3)^3\left((4a+t_3)^3-4(2a+t_3)^3\right) \qquad -4a\le s_3\le -2a \text{ and } -2a\le t_3\le 0$$

$$f(s_3,t_3)=\frac{1}{9216a^8}(4a+s_3)^3\left((4a-t_3)^3-4(2a-t_3)^3\right) \qquad -4a\le s_3\le -2a \text{ and } 0\le t_3\le 2a$$

$$f(s_3,t_3)=\frac{1}{9216a^8}(4a+s_3)^3(4a-t_3)^3 \qquad -4a\le s_3\le -2a \text{ and } 2a\le t_3\le 4a$$

$$f(s_3,t_3)=\frac{1}{9216a^8}\left((4a+s_3)^3-4(2a+t_3)^3\right)(4a-t_3)^3 \qquad -2a\le s_3\le 0 \text{ and } -4a\le t_3\le -2a$$

$$f(s_3,t_3)=\frac{1}{9216a^8}\left((4a+s_3)^3-4(2a+t_3)^3\right)\left((4a+t_3)^3-4(2a+t_3)^3\right)$$

$$-2a\le s_3\le 0 \text{ and } -2a\le t_3\le 0$$

$$f(s_3,t_3)=\frac{1}{9216a^8}\left((4a+s_3)^3-4(2a+t_3)^3\right)\left((4a-t_3)^3-4(2a-t_3)^3\right)$$

$$-2a\le s_3\le 0 \text{ and } 0\le t_3\le 2a$$

$$f(s_3,t_3)=\frac{1}{9216a^8}\left((4a+s_3)^3-4(2a+s_3)^3\right)(4a-t_3)^3 \qquad -2a\le s_3\le 0 \text{ and } 2a\le t_3\le 4a$$

$$f(s_3,t_3)=\frac{1}{9216a^8}\left((4a-s_3)^3-4(2a-t_3)^3\right)(4a+t_3)^3 \qquad 0\le s_3\le 2a \text{ and } -4a\le t_3\le -2a$$

$$f(s_3,t_3)=\frac{1}{9216a^8}\left((4a-s_3)^3-4(2a-t_3)^3\right)\left((4a+t_3)^3-4(2a+t_3)^3\right)$$

$$0\le s_3\le 2a \text{ and } -2a\le t_3\le 0$$

$$f(s_3,t_3)=\frac{1}{9216a^8}\left((4a-s_3)^3-4(2a-t_3)^3\right)\left((4a-t_3)^3-4(2a-t_3)^3\right)$$

$$0\le s_3\le 2a \text{ and } 0\le t_3\le 2a$$

$$f(s_3,t_3)=\frac{1}{9216a^8}\left((4a-s_3)^3-4(2a-s_3)^3\right)(4a-t_3)^3 \qquad 0\le s_3\le 2a \text{ and } 2a\le t_3\le 4a$$

$$f(s_3,t_3)=\frac{1}{9216a^8}(4a-s_3)^3(4a+t_3)^3 \qquad 2a\le s_3\le 4a \text{ and } -4a\le t_3\le -2a$$

$$f(s_3,t_3)=\frac{1}{9216a^8}(4a-s_3)^3\left((4a+t_3)^3-4(2a+t_3)^3\right) \qquad 2a\le s_3\le 4a \text{ and } -2a\le t_3\le 0$$

$$f(s_3,t_3)=\frac{1}{9216a^8}(4a-s_3)^3\left((4a-t_3)^3-4(2a-t_3)^3\right) \qquad 2a\le s_3\le 4a \text{ and } 0\le t_3\le -2a$$

$$f(s_3,t_3)=\frac{1}{9216a^8}(4a-s_3)^3(4a-t_3)^3 \qquad 2a\le s_3\le 4a \text{ and } 2a\le t_3\le 4a$$

Interpolation and extrapolation (linear and cubic spline)

$$\hat{y}^P=y_0^P+\hat{v}^P=\sum_{j=i}^{i+1}\hat{a}_j S_{j-i+1}^{j}\left(s_P^j\right) \qquad \forall P\in A_i \;\; i=1,n-1 \qquad s_P^j=\left(t^P-t^j\right)/\Delta t$$

$$\hat{y}^p=y_0^p+\hat{v}^p=\sum_{j=i-1}^{i+2}\hat{a}_j S_{j-i+2}^{j}\left(s_p^j\right) \qquad \forall P\in A_i \;\; i=2,n-2$$

(bilinear and bi-cubic spline)

$$\hat{z}^p=z_0^p+\hat{v}^p=\sum_{k'=i}^{i+1}\sum_{k''=j}^{j+1}\hat{a}_{(k''-1)l+k'}S_{2(k''-j)+k'-i+1}^{(k''-1)l+k'}\left(s_p^{k'},t_p^{k''}\right) \qquad \forall P\in A_i \quad i=1,l-1 \text{ and } j=1,h-1 \;\; (lh=n)$$

$$s_p^{k'} = \left(x^p - x^{k'}\right)/\Delta x\,,\; t_p^{k''} = \left(y^p - y^{k''}\right)/\Delta y$$

$$\hat{z}^p = z_0^p + \hat{v}^p = \sum_{k'=i-1}^{i+2} \sum_{k''=j-1}^{j+2} \hat{a}_{(k''-1)l+k'} S_{4(k''-j+1)+k'-i+2}^{(k''-1)l+k'} \left(s_p^{k'}, t_p^{k''}\right) \qquad \forall P \in A_i$$

$$i = l-3 \text{ and } j = h-3 \;\; (lh = n)$$

COVARIANCE ESTIMATION

Empirical estimation:

- general variance: $\sigma^2 = \dfrac{1}{n-1}\sum_{i=1}^{n} v_i^2$ with: $v = x - \overline{x}$

- 1D auto-covariance (in the interval ΔT):

$$\gamma(\Delta T_k) = \frac{1}{n}\sum_{i=1}^{n} v_i \frac{1}{n_i}\sum_{j=1}^{n_i} v_j$$

where: $T_{k-1} < \left|t_i - t_{ju}\right| \le T_k$ and $\Delta T_k = T_k - T_{k-1}$

- 2D (or 3D) auto-covariance (in the neighborhood ΔP):

$$\gamma(\Delta P_k) = \frac{1}{n}\sum_{i=1}^{n} v_i \frac{1}{n_i}\sum_{j=1}^{n_j} v_j$$

where: $P_{k-1} < \left\|Q_i - Q_j\right\| \le P_k$ and $\Delta P_k = P_k - P_{k-1}$

- correlation coefficient: $\rho = \gamma / \sigma^2$

- general covariance:

$$\gamma_{xy} = \frac{1}{l}\sum_{i=1}^{l} v_i u_i$$

with: $v = x - \overline{x}$ and $u = y - \overline{y}$

- 1D cross-covariance (in the interval Δt):

$$\gamma_{xy}(\Delta T_k) = \frac{1}{2}\left(\frac{1}{n}\sum_{i=1}^{n} v_i \frac{1}{m_i}\sum_{j=1}^{m_j} u_j + \frac{1}{m}\sum_{i=1}^{m} u_i \frac{1}{n_i}\sum_{j=1}^{n_i} v_j\right)$$

where $T_{k-1} < \left|t_i - t_j\right| \le T_k$ and $\Delta T_k = T_k - T_{k-1}$

- 2D (or 3D) cross-covariance (in the neighborhood ΔP):

$$\gamma_{xy}(\Delta P_k) = \frac{1}{2}\left(\frac{1}{n}\sum_{i=1}^{n} v_i \frac{1}{m_i}\sum_{j=1}^{m_j} u_j + \frac{1}{m}\sum_{i=1}^{m} u_i \frac{1}{n_i}\sum_{j=1}^{n_i} v_j\right)$$

where: $P_{k-1} < \left\|Q_i - Q_j\right\| \le P_k$ and $\Delta P_k = P_k - P_{k-1}$

- correlation coefficient: $\rho_{xy} = \gamma_{xy} / \sigma_x \sigma_y$

Empirical optimization of the spacing: $\sigma_n^2 = \sigma^2 - \sigma_s^2 \cong \sigma^2 - \gamma_1 = \tilde{\sigma}_n^2 \qquad \Rightarrow$

$$max\,\gamma_1 = min\,\tilde{\sigma}_n^2 \cong min\,\sigma_n^2$$ according to different values of Δt or ΔP.

Theoretical models:

$\gamma = ae^{-b|\xi|}\cos(c\xi)$ (1D only)

$\gamma = ae^{-b\xi^2}\cos(c\xi)$ (1D only)

$$\gamma = ae^{-b|\xi|}\frac{\sin(c\xi)}{c\xi}$$

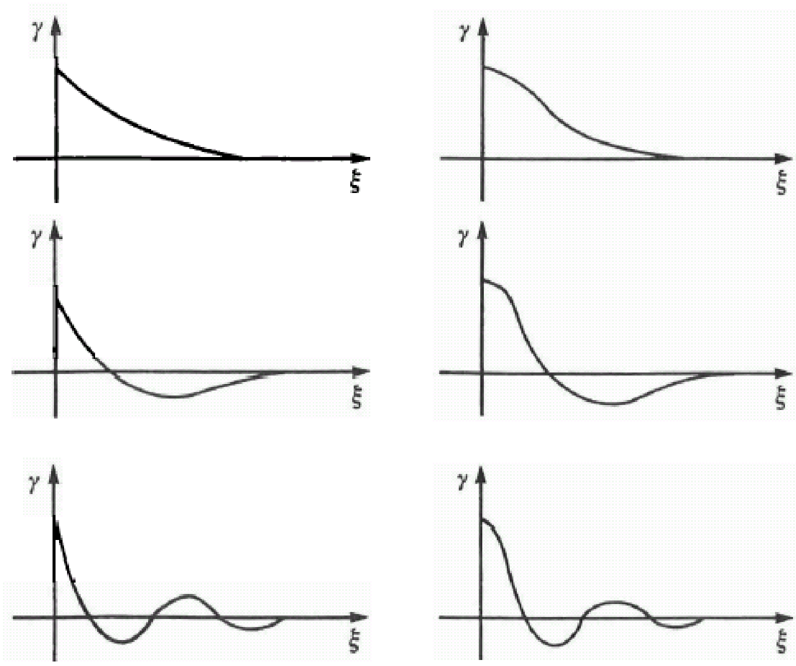
$$\gamma = ae^{-b\xi^2}\frac{\sin(c\xi)}{c\xi}$$

$\gamma = ae^{-b|\xi|}J_0(c\xi)$ (1D and 2D)

$\gamma = ae^{-b\xi^2}J_0(c\xi)$ (1D and 2D)

$$\gamma = 2ae^{-b|\xi|}\frac{J_1(c\xi)}{c\xi}$$

$$\gamma = 2ae^{-b\xi^2}\frac{J_1(c\xi)}{c\xi}$$



Examples of theoretical models

Note: A linear combination (with positive coefficients, like the sum and the weighted sum), as well as the product and the convolution in itself, preserves the covariance function properties (e.g. positive Eigen-values in the corresponding matrices) and supplies additional models.

Note: Cross-covariance functions are less important than the auto-covariance ones and they are only used to better refine a previous optimal solution.

Finite covariance functions:

$$S_1(t)=\int_{t-a}^{a}(a^2-x^2)(a^2-(t-x)^2)dx=\frac{16a^5}{15}-\frac{4a^3t^2}{3}+\frac{2a^2t^3}{3}-\frac{t^5}{30} \qquad \text{with } t\le 2a$$

$$S_1(t)=0 \qquad \text{with } t\ge 2a$$

$$S_2(r)=\int_{r/2}^{a}dx\int_{0}^{\sqrt{a^2-x^2}}dy\,(a^2-x^2-y^2)(a^2-(r-x)^2-y^2)= \qquad \text{with } r\le 2a$$

$$=\frac{4}{15}\int_{r/2}^{a}(8x^4-20rx^3+(10r^2-16a^2)x^2+20a^2rx-10a^2r^2+8a^4)\sqrt{a^2-x^2}\,dx=$$

$$=5a^6\pi-\frac{15a^4r^2\pi}{2}+(5a^4r+\frac{20a^2r^3}{3}-\frac{5r^5}{12})\frac{\sqrt{a^2-r^2}}{4}+(15a^4r^2-10a^6)\,arcsin\frac{r}{2a}$$

$$S_2(r)=0 \qquad \text{with } r\ge 2a$$

<u>Note:</u> Finite covariance functions supply sparse matrices, which are easier to treat numerically.

<u>Note:</u> 3D finite covariance functions are supplied by the orthogonal product: 1D × 2D (or 1D × 1D × 1D), being the results quasi isotropic only.

Space-temporal problems: $\gamma(\Delta P,\Delta T)=\gamma(\Delta P)\bullet\gamma(\Delta T)$ $\Rightarrow$

$C(\Delta P,\Delta T)=C(\Delta P)\otimes C(\Delta T)$ $\otimes$ being the Kronecker product.

COLLOCATION (FILTERING AND PREDICTION)

- Functional and stochastic models:

$v=s+n$ where: s signal and n noise

$C_{vv}=C_{ss}+\sigma_n{}^2 I$ where: $\sigma_v^2=\sigma_s^2+\sigma_n^2$

- Hybrid norm:

$$w^T\,C_{ww}{}^{-1}w\;+\;\lambda^T\left(Aw-v\right)=\;min$$

where: $w=\begin{vmatrix} n \\ s \\ t \end{vmatrix}$ t being the predicted signal

- Estimates:

$\hat{s}=C_{ss}\,C_{vv}{}^{-1}v$ $C_{\hat{s}\hat{s}}=C_{ss}\,C_{vv}{}^{-1}\,C_{ss}$ $C_{ee}=C_{ss}-C_{ss}\,C_{vv}{}^{-1}\,C_{ss}$

$\hat{n}=\sigma_n{}^2\,C_{vv}{}^{-1}v$ $C_{\hat{n}\hat{n}}=\sigma_n^4\,C_{vv}{}^{-1}$

$\hat{t}=C_{ts}\,C_{vv}{}^{-1}v$

or better:

$\hat{n} = \sigma_n^{\ 2}\ C_{vv}^{\ -1} v$ $\qquad C_{\hat{n}\hat{n}} = \sigma_n^4\ C_{vv}^{\ -1}$

$\hat{s} = v - \hat{n}$ $\qquad C_{\hat{s}\hat{s}} = C_{ss} - \sigma_n^2\ \left(I - \sigma_n^2\ C_{vv}^{\ -1} \right)$ $\qquad C_{ee} = \sigma_n^2 \left(I - \sigma_n^2\ C_{vv}^{\ -1} \right)$

$\hat{t} = C_{ts}\ z$ $\qquad$ where: $z = C_{vv}^{\ -1}\ v$

GENERALIZED LEAST SQUARES

- Functional model and hybrid norm:

$y = Ax + Bs$ $\qquad y_0 = A\hat{x} + B\hat{s} - \hat{n}$

$$1/2\begin{bmatrix}\hat{s}^T & \hat{n}^T\end{bmatrix}\begin{bmatrix}C_{ss}^{-1} & 0 \\ 0 & P/\sigma_n^2\end{bmatrix}\begin{bmatrix}\hat{s} \\ \hat{n}\end{bmatrix} + \lambda'\left(A\hat{x} + B\hat{s} - \hat{n} - y_0\right) = min$$

$y = B\,s$ $\qquad y_0 = B\hat{s} - \hat{n}$

where: $s^T = \begin{bmatrix}x^T & s^T\end{bmatrix}$ $\qquad$ and $\qquad C_{ss} = \begin{bmatrix}hI & 0 \\ 0 & C_{ss}\end{bmatrix}$

$$1/2\begin{bmatrix}\hat{s}^T & \hat{n}^T\end{bmatrix}\begin{bmatrix}C_{ss}^{-1} & 0 \\ 0 & P/\sigma_n^2\end{bmatrix}\begin{bmatrix}\hat{s} \\ \hat{n}\end{bmatrix} + \lambda^T\left(B\hat{s} - \hat{n} - y_0\right) = min$$

- Estimates:

$\hat{s} = C_{ss}B^T\left(BC_{ss}B^T + \sigma_n^2 P^{-1}\right)^{-1} y_0$ $\qquad C_{ee} = C_{ss}\left(I - B^T\left(BC_{ss}B^T + \sigma_n^2 P^{-1}\right)^{-1}\right)BC_{ss}$

$\hat{n} = y_0 - B\hat{s}$ $\qquad C_{\hat{n}\hat{n}} = \sigma_n^4 P^{-1}\left(BC_{ss}B^T + \sigma_n^2 P^{-1}\right)^{-1}$

$\hat{t} = C_{ts}z$ $\qquad$ being: $z = B^T\left(BC_{ss}B^T + \sigma_n^2 P^{-1}\right)^{-1} y_0$

or better:

$$\hat{s} = \left(B^T PB\right)^{-1} B^T Py_0 - \sigma_n^2\left(B^T PBC_{ss}B^T PB + \sigma_n^2 B^T PB\right)^{-1} B^T Py_0$$

$$C_{ee} = \sigma_n^2\left(B^T PB\right)^{-1} - \sigma_n^4\left(B^T PBC_{ss}B^T PB + \sigma_n^2 B^T PB\right)^{-1}$$

$\hat{n} = y_0 - B\hat{s}$ $\qquad C_{\hat{n}\hat{n}} = \sigma_n^2 P^{-1} - BC_{ee}B^T$

$\hat{t} = C_{ts}z$ $\qquad$ where: $z = B^T PB\left(B^T PBC_{ss}B^T PB + \sigma_n^2 B^T PB\right)^{-1} B^T P y_0$

because: $\left(Q \pm RST\right)^{-1} = Q^{-1} \mp Q^{-1}R\left(S^{-1} \pm TQ^{-1}R\right)^{-1} TQ^{-1}$

and $Q^{-1}\left(Q^{-1} \pm S\right)^{-1} Q^{-1} = \left(Q \pm QSQ\right)^{-1}$

$$\hat{\sigma}_n^2 = \frac{n^T Pn}{m - n + Tr\left(\sigma_n^2 P^{1/2} B\left(B^T PBC_{ss}B^T PB + \sigma_n^2 B^T PB\right)^{-1} B^T P^{1/2}\right)}$$

KRIGING

- Semi-variogram: $\omega(\Delta_k) = \dfrac{1}{2n_k}\sum_{i=1}^{n_k}\left(z(x_i+\Delta_k)-z(x_i)\right)^2 \quad \forall\Delta_k$

Note: $\gamma(\Delta_k) = \sigma^2 - \omega(\Delta_k)$ where:

$$\omega(\Delta_k) = \frac{1}{2n_k}\sum_{i=1}^{n_k} z(x_i+\Delta_k)^2 + \frac{1}{2n_k}\sum_{i=1}^{n_k} z(x_i)^2 - \frac{1}{n_k}\sum_{i=1}^{n_k} z(x_i+\Delta_k)\,z(x_i) = \sigma^2 - \gamma(\Delta_k) \quad \forall\Delta_k$$

- Simple Kriging :

$$\hat{z}(x_0) = \gamma^T(x_0, x_i)\,\Gamma^{-1}(x_i, x_j)\,z(x_j) = w^T(x_i)\,z(x_j)$$

$$\hat{\sigma}^2(x_0) = \sigma^2(x_0) - \gamma^T(x_0, x_i)\,\Gamma^{-1}(x_i, x_j)\,\gamma(x_j, x_0)$$

- Ordinary Kriging:

$$\hat{z}(x_0) = \begin{bmatrix}\omega^T(x_0, x_i) & 1\end{bmatrix}\begin{bmatrix}\Omega(x_i, x_j) & i \\ i^T & 0\end{bmatrix}^{-1}\begin{bmatrix} z(x_j) \\ 0\end{bmatrix} = \begin{bmatrix} w^T(x_i) & \lambda\end{bmatrix}\begin{bmatrix} z(x_j) \\ 0\end{bmatrix}$$

$$\hat{\sigma}^2(x_0) = \begin{bmatrix}\omega^T(x_0, x_i) & 1\end{bmatrix}\begin{bmatrix}\Omega(x_i, x_j) & i \\ i^T & 0\end{bmatrix}^{-1}\begin{bmatrix}\omega(x_j, x_0) \\ 1\end{bmatrix}$$

- Ordinary Co-Kriging:

$$\hat{z}_1(x_0) = \begin{bmatrix}\omega_{11}(x_0, x_j) & \omega_{12}(x_0, x_l) & 1 & 0\end{bmatrix}\begin{bmatrix}\Omega_{11}(x_i, x_j) & \Omega_{12}(x_i, x_l) & i & 0 \\ \Omega_{12}^T(x_i, x_j) & \Omega_{22}(x_i, x_l) & 0 & j \\ i^T & 0 & 0 & 0 \\ 0 & j^T & 0 & 0\end{bmatrix}^{-1}\begin{bmatrix} z_1(x_j) \\ z_2(x_l) \\ 0 \\ 0\end{bmatrix} =$$

$$= \begin{bmatrix} w_{11}(x_i) & w_{12}(x_i) & \lambda_1 & \mu_1\end{bmatrix}\begin{bmatrix} z_1(x_j) \\ z_2(x_l) \\ 0 \\ 0\end{bmatrix}$$

$$\hat{z}_2(x_0) = \begin{bmatrix}\omega_{21}(x_0, x_j) & \omega_{22}(x_0, x_l) & 0 & 1\end{bmatrix}\begin{bmatrix}\Omega_{11}(x_k, x_j) & \Omega_{12}(x_k, x_l) & 0 & i \\ \Omega_{12}^T(x_k, x_j) & \Omega_{22}(x_k, x_l) & j & 0 \\ 0 & j^T & 0 & 0 \\ i^T & 0 & 0 & 0\end{bmatrix}^{-1}\begin{bmatrix} z_1(x_j) \\ z_2(x_l) \\ 0 \\ 0\end{bmatrix} =$$

$$= \begin{bmatrix} w_{21}(x_k) & w_{22}(x_k) & \lambda_2 & \mu_2\end{bmatrix}\begin{bmatrix} z_1(x_j) \\ z_2(x_l) \\ 0 \\ 0\end{bmatrix}$$

$$\hat{\sigma}_1^2(x_0) = \begin{bmatrix} \omega_{11}(x_0, x_j) & \omega_{12}(x_0, x_l) & 1 & 0 \end{bmatrix} \begin{bmatrix} \Omega_{11}(x_k, x_j) & \Omega_{12}(x_k, x_l) & i & 0 \\ \Omega_{12}^T(x_k, x_j) & \Omega_{22}(x_k, x_l) & 0 & j \\ i^T & 0 & 0 & 0 \\ 0 & j^T & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \omega_{11}(x_j, x_0) \\ \omega_{12}(x_l, x_0) \\ 1 \\ 0 \end{bmatrix}$$

$$\hat{\sigma}_2^2(x_0) = \begin{bmatrix} \omega_{21}(x_0, x_j) & \omega_{22}(x_0, x_l) & 0 & 1 \end{bmatrix} \begin{bmatrix} \Omega_{11}(x_k, x_j) & \Omega_{12}(x_k, x_l) & 0 & i \\ \Omega_{12}^T(x_k, x_j) & \Omega_{22}(x_k, x_l) & j & 0 \\ 0 & j^T & 0 & 0 \\ i^T & 0 & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \omega_{21}(x_j, x_0) \\ \omega_{22}(x_l, x_0) \\ 0 \\ 1 \end{bmatrix}$$

- Universal Kriging:

$$\hat{z}(x_0) = \begin{bmatrix} \omega^T(x_0, x_i) & 1 & f^T(x_0) \end{bmatrix} \begin{bmatrix} \Omega(x_i, x_j) & i & F_k(x_i) \\ i^T & 0 & 0 \\ F^T(x_i) & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} z(x_j) \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} w^T(x_i) & \lambda & \mu \end{bmatrix} \begin{bmatrix} z(x_j) \\ 0 \\ 0 \end{bmatrix}$$

$$\hat{\sigma}^2(x_0) = \begin{bmatrix} \omega^T(x_0, x_i) & 1 & f^T(x_0) \end{bmatrix} \begin{bmatrix} \Omega(x_i, x_j) & i & F(x_i) \\ i^T & 0 & 0 \\ F^T(x_i) & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \omega(x_j, x_0) \\ 1 \\ f(x_0) \end{bmatrix}$$

- Universal Co-Kriging: ...

  ...

Note: Simple Kriging is equivalent to Collocation (filtering and prediction).

Note: Collocation (filtering and prediction) is a particular case of the Generalized least squares (setting: $B = 0$):

$$\hat{s} = y_0 - \hat{n}_0 \qquad C_{ee} = \sigma_n^2 I - C_{\hat{n}\hat{n}}$$

$$\hat{n} = \sigma_n^2 \left(C_{ss} + \sigma_n^2 I\right)^{-1} y_0 \qquad C_{\hat{n}\hat{n}} = \sigma_n^4 \left(C_{ss} + \sigma_n^2 I\right)^{-1}$$

furthermore the same problem could be reduced to classical least squares (setting: $\sigma_n^2 = 0$ and $A = B$, and substituting $\hat{s}$ with $\hat{x}$ and $\hat{n}$ with $\hat{v}$):

$$\hat{x} = -\left(A^T P A\right)^{-1} A^T P y_0 \qquad C_{\hat{x}\hat{x}} = \sigma_0^2 \left(A^T P A\right)^{-1}$$

$$\hat{v} = y_0 - A\hat{x} = y_0 - \hat{y} \qquad C_{\hat{v}\hat{v}} = \sigma_0^2 P^{-1} - A C_{\hat{x}\hat{x}} A^T = \sigma_0^2 P^{-1} - C_{\hat{y}\hat{y}}$$

Note: Interpolation and approximation theory, together with multivariate analysis methods, prove the power and the richness of the statistical methodologies and procedures. Indeed these algorithms and techniques are able to explore the wide space of the data, especially if they are spatially referenced (whether time dependent or not), supplying deterministic, semi-deterministic or stochastic models and completing the information with metadata. The last ones are produced by means of a refined statistical analysis, exploiting tools of either parametric or distribution-free statistical inference. Both estimated data and related metadata can now be easily achieved by the capabilities of computational statistics and its numerical methods. A large number of examples proves these assertions, such as optimal sampling, network adjustment and field reconstruction (e.g. line, surface, 3D model, etc.) can very clearly show and positively confirm.

PART V – NUMERICAL METHODS

DIRECT METHODS

Gauss' elimination

$$c'_{jk} = c_{jk} - c_{ik} \cdot \frac{c_{ji}}{c_{ii}} \qquad \forall j,k$$

$$d'_{j} = -\left(d_{j} - d_{i} \frac{c_{ji}}{c_{ii}} \right) \qquad \forall j$$

$$x_i = d_i - \sum_{k=i+1}^{n} c_{ik} x_k \qquad \forall i$$

Cholesky factorization

$$T^T T = C$$

$$c_{11} = t_{11} \cdot t_{11}$$

$$c_{1j} = t_{11} \cdot t_{1j} \qquad (j > 1)$$

$$c_{ii} = t_{ii} \cdot t_{ii} + \sum_{k=1}^{i-1} t_{ki} \cdot t_{ki} \qquad (i \neq 1)$$

$$c_{ij} = t_{ii} \cdot t_{ij} + \sum_{k=1}^{i-1} t_{ki} \cdot t_{kj} \qquad (j > 1)$$

$$t_{11} = \sqrt{c_{11}}$$

$$t_{1j} = \frac{c_{1j}}{t_{11}} \qquad (j > 1)$$

$$t_{ii} = \sqrt{c_{ii} - \sum_{k=1}^{i-1} t_{ki}^2} \qquad (i \neq 1)$$

$$t_{ij} = \frac{\left(c_{ij} - \sum_{k=1}^{i-1} t_{ki} \cdot t_{kj} \right)}{t_{ii}} \qquad (j > 1)$$

Forward and backward substitution

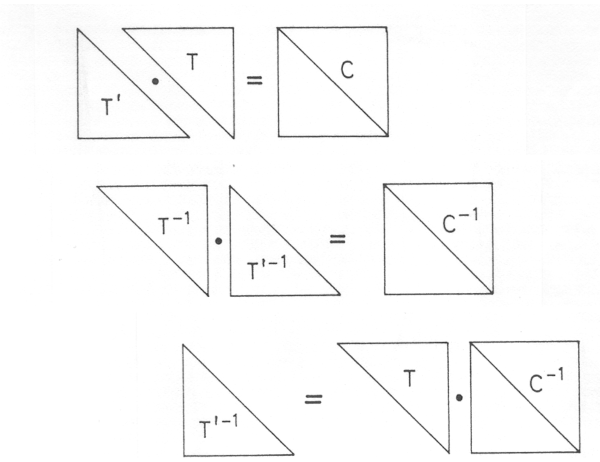
$$Cx + d = 0 \qquad T^T T x = -d \qquad \begin{aligned} T^T y &= -d \\ T x &= y \end{aligned}$$

$$t_{11} y_1 = -d_1 \qquad y_1 = -\frac{d_1}{t_{11}}$$

$$t_{12}\, y_1 + t_{22}\, y_2 = -d_2 \qquad y_2 = -\frac{d_2 + t_{12} y_1}{t_{22}}$$

$$y_1 = -\frac{x_1}{t_{11}} \qquad x_n = \frac{y_n}{t_{nn}}$$

$$y_i = -\frac{d_i + \sum_{k=1}^{i-1} t_{ki} y_k}{t_{ii}} \qquad (i \neq 1) \qquad\qquad {}_i = \frac{y_i - \sum_{k=i+1}^{n} t_{ik} x_k}{t_{ii}} \qquad (i \neq n)$$



Cholesky factors

Inverse matrix

Inversion of the normal matrix

$$C^{-1} = \left(T^T T\right)^{-1} = T^{-1}\left(T^T\right)^{-1}$$

$$T\,C^{-1} = \left(T^T\right)^{-1} \qquad\qquad : T\,col_i\left(C^{-1}\right) = col_i\left(\left(T^T\right)^{-1}\right)$$

$$\sum_{k=i}^{n} t_{ik}\gamma_{ik} = \frac{1}{t_{ii}} \qquad\qquad \sum_{k=i}^{n} t_{ik}\gamma_{jk} = 0 \qquad (j > i)$$

$$\gamma_{nn} = \frac{1}{t_{nn}^2}$$

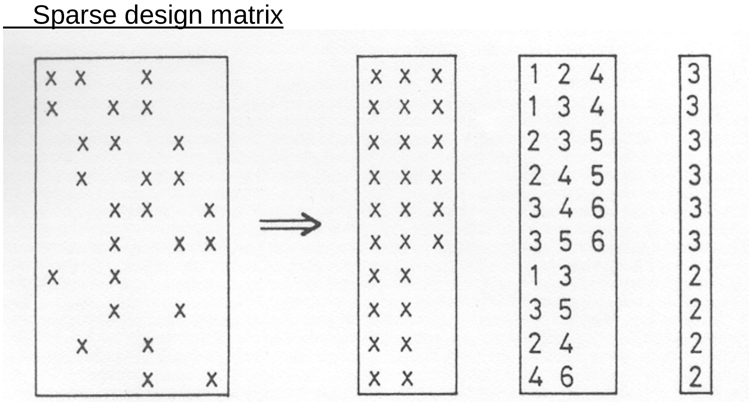
$$\gamma_{ij} = -\frac{\sum_{k=i+1}^{n} t_{ik}\gamma_{\substack{kj\ (se\ j \geq k)\\ jk\ (se\ j<k)}}}{t_{ii}} = -\frac{\sum_{k=i+1}^{j} t_{ik}\gamma_{kj} + \sum_{k=j+1}^{n} t_{ik}\gamma_{jk}}{t_{ii}} \qquad (j > i)$$

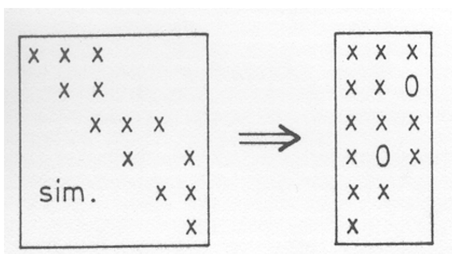
$$\gamma_{ii} = \frac{\frac{1}{t_{ii}} - \sum_{k=i+1}^{n} t_{ik}\gamma_{ik}}{t_{ii}} \qquad (i \neq n)$$

Sparse matrices

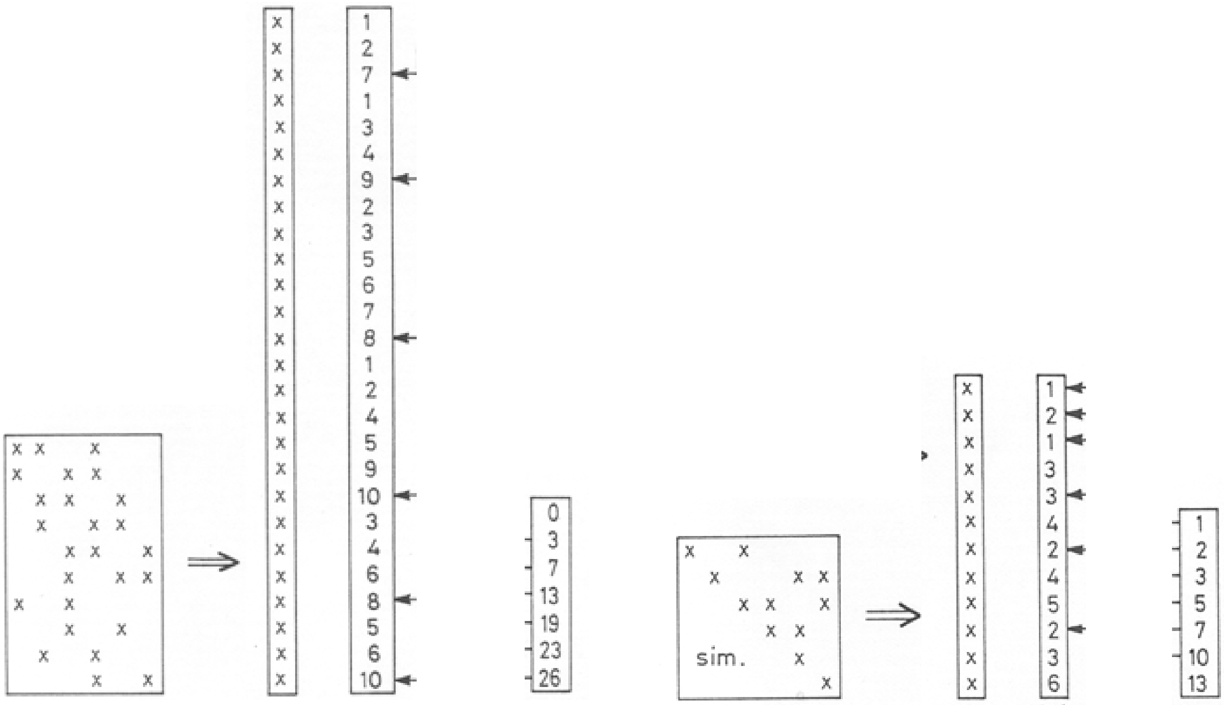
Sparse design matrix

Banded normal matrix

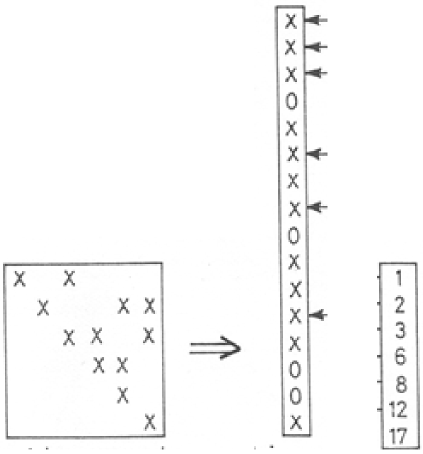




Sparse normal matrix



Profile normal matrix



Variance propagation

$$y = \sum_{i=1}^{l} a_i x_i \qquad\qquad z = \sum_{i=1}^{l} b_i x_i$$

$$\sigma_y^2 = \sum_{i=1}^{l} a_i^2 \sigma_{x_i}^2 + 2\sum_{i=1}^{l} \sum_{j=i+1}^{l} a_i a_j \sigma_{x_i x_j} = \sum_{i=1}^{l}\left(2\sum_{j=i}^{l} a_i a_j \sigma_{x_i x_j} - a_i^2 \sigma_{x_i}^2\right)$$

$$\sigma_z^2 = \ldots$$

$$\sigma_{yz} = \sum_{i=1}^{l} a_i b_i \sigma_{x_i}^2 + \sum_{i=1}^{l} \sum_{j=i+1}^{l} \left(a_i b_j + a_j b_i\right)\sigma_{x_i x_j} = \sum_{i=1}^{l}\left(\sum_{j=i}^{l} \left(a_i b_j + a_j b_i\right)\sigma_{x_i x_j} - a_i b_i \sigma_{x_i}^2\right)$$

$$r_{yz} = \sigma_{yx} \big/ \sigma_y \sigma_z$$

ITERATIVE METHODS

Conjugate Gradient method

$$p_1 = r_1 = -\left(Cx_o + d\right)$$

$$\alpha_i = \frac{r_i^T\, r_i}{p_i^T\, C\, p_i}$$

$$x_i = x_{i-1} + \alpha_i\, p_i$$

$$r_{i+1} = r_i - \alpha_i\, C\, p_i \qquad\qquad i = 1,2,........,\tilde{n}$$

$$\beta_i = \frac{r_{i+1}^T\, r_{i+1}}{r_i^T\, r_i}$$

$$p_{i+1} = r_{i+1} + \beta_i\, p_i$$

$$\tilde{x} = x_{\tilde{n}}$$

Preconditioning and ICCG methods

Incomplete Cholesky factorization

$$w_{ii} = \sqrt{c_{ii} - \sum_{k=1}^{i-1} w_{ki}^2}$$

$$w_{ij} = \frac{\left(c_{ij} - \sum_{k=1}^{i-1} w_{ki}\, w_{kj}\right)}{w_{ii}} \qquad \left(j > i\right) \qquad \text{if } c_{ij} \neq 0$$

$$w_{ij} = 0 \qquad \left(j > i\right) \qquad \text{if } c_{ij} = 0$$

ICCG method

$$r_1 = -\left(C\, x_o + d\right)$$

$$p_1 = \left(W^T W\right)^{-1} r_1$$

$$\alpha_i = \frac{r_i^T \left(W^T W\right)^{-1} r_i}{p_i^T\, C\, p_i}$$

$$x_i = x_{i-1} + \alpha_i\, p_i$$

$$r_{i+1} = r_i - \alpha_i\, C\, p_i \qquad i = 1,2,.......,\tilde{n}$$

$$\beta_i = \frac{r_{i+1}^T \left(W^T W\right)^{-1} r_{i+1}}{r_i^T \left(W^T W\right) r_i}$$

$$p_{i+1} = \left(W^T W\right)^{-1} r_{i+1} + \beta_i p_i$$

$$\tilde{x} = x_{\tilde{n}}$$

Approximate inversion

$$\gamma_{ii} = \frac{\dfrac{1}{w_{ii}} - \sum_{k=1+1}^{n} w_{ik}\gamma_{ik}}{w_{ii}}$$

$$\gamma_{ij} = \frac{-\left(\sum_{k=i+1}^{j} w_{ik}\gamma_{kj} + \sum_{k=j+1}^{n} w_{ik}\gamma_{jk}\right)}{w_{ii}} \qquad (j > i) \qquad \text{if } c_{ij} \neq 0$$

$$\gamma_{ij} = 0 \qquad (j > i) \qquad \text{if } c_{ij} = 0$$

REGULAR STRUCTURES

Toeplitz matrix

$$W = \lfloor \omega_{ij} \rfloor = \lfloor \varphi_{|i-j|} \rfloor \qquad i, j = 1, n$$

Properties of its inverse matrix

$\omega_{ij} = \omega_{ji} \quad \forall i, j$ (Hermite Symmetry)

$\omega_{ij} = \omega_{n-j+1\ n-i+1} \quad \forall i, j$ (Wise Per-symmetry)

Kronecker decomposition

$$(U \otimes V) = \lfloor u_{ij} V \rfloor = \lfloor u_{ij} [v_{kl}] \rfloor \qquad i, j = 1, m \quad k, l = 1, n$$

Properties of its inverse matrix and related system

$$(U \otimes V)^{-1} = U^{-1} \otimes V^{-1}$$

$(U \otimes V)x = \vartheta$ i.e.: $x = (U \otimes V)^{-1}\vartheta = U^{-1} \otimes V^{-1}\vartheta$

$$y_i = V^{-1}\vartheta_i \qquad i = 1, m$$
$$x_j = U^{-1} z_j \qquad j = 1, n$$

where: $z_j = \lfloor y_{ij} \rfloor \quad i = 1, m \quad j = 1, n$.

and in matrix form: $Y = V^{-1}\Theta$ and $X = U^{-1}Z$

where: $\Theta = [\vartheta_i]$ $\quad Y = [y_i]$ $\quad i = 1, m$

$Z = [z_j]$ $\quad X = [x_j]$ $\quad j = 1, n$

being: $Z = Y^T$

Trench algorithm

$$f_l^{(l)} = \left(b_l - \sum_{i=1}^{l-1} f_i^{(l-1)} b_{l-i} \Big/ e^{(l-1)} \right)$$

$$f_i^{(l)} = f_i^{(l-1)} - f_l^{(l)} \cdot f_{l-i}^{(l-1)} \qquad i = 1, l-1 \qquad ; \qquad l = 2, n-1$$

$$e^{(l)} = \left(1 - f_l^{2(l)}\right) \cdot e^{(l-1)}$$

where: $e^{(1)} = 1$ $\qquad f_1^{(1)} = b_1$

System solution:

$$\begin{pmatrix} 1 & b^T \\ b & C \end{pmatrix} \begin{pmatrix} y \\ x \end{pmatrix} = \begin{pmatrix} \eta \\ \xi \end{pmatrix} \qquad \begin{pmatrix} y \\ x \end{pmatrix} = \begin{pmatrix} \alpha\left(\eta - b^T x_0\right) \\ x_0 - C^{-1} by \end{pmatrix} = \begin{pmatrix} \left(\eta - b^T x_0\right)/e \\ x_0 - fy \end{pmatrix}$$

where: $p_{11} = 1$ and $y = \eta$

Inverse matrix:

$$\psi_{11} = \alpha = 1/e$$

$$\psi_{1,i} = \beta_{i-1} = -f_{i-1}/e \qquad i = 2, n$$

$$\psi_{ij} = \gamma_{i-1,j-1} = \psi_{i-1,j-1} + \left(f_{i-1} f_{j-1} - f_{n-i+1} f_{n-j+1}\right)/e \qquad i, j = 2, n$$

with its diagonal elements: $\omega_1 = 1/e$ $\qquad \omega_i = \omega_{i-1} + \left(f_{i-1}^2 - f_{n-i+1}^2\right)/e$ $\qquad i = 2, n$

SEQUENTIAL ALGORITHMS

Householder transformation

Initialization: $A^{(1)} = A$ $\qquad b^{(1)} = b$

Assignments: $\alpha^{(k)} = sign\left(a_{kk}^{(k)}\right)\sqrt{\sum_{i=k}^{m}\left(a_{ik}^{(k)}\right)^2}$ $\qquad \beta^{(k)} = \dfrac{1}{\alpha^{(k)}\left(\alpha^{(k)} + \alpha_{kk}^{(k)}\right)}$

$$\psi_j^{(k)} = \beta^{(k)}\left(\left(\alpha^{(k)} + \alpha_{kk}^{(k)}\right) \alpha_{kj}^{(k)} + \sum_{i=k+1}^{m} \alpha_{ik}^{(k)} \alpha_{ij}^{(k)} \right) \qquad j > k$$

$$\xi^{(k)} = \beta^{(k)}\left(\left(\alpha^{(k)} + a_{kk}^{(k)}\right) b_k^{(k)} + \sum_{i=k+1}^{m} a_{ik}^{(k)} b_i^{(k)} \right)$$

The algorithm:

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} \qquad \forall\, i, \quad j < k$$

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} \qquad i < k,\ j \ge k$$

$$a_{kk}^{(k+1)} = -\alpha^{(k)}$$

$$a_{ik}^{(k+1)} = 0 \qquad i > k$$
$$a_{kj}^{(k+1)} = a_{kj}^{(k)} - \left(\alpha^{(k)} + a_{kk}^{(k)}\right)\psi_j^{(k)} \qquad j > k$$

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - a_{ik}^{(k)}\psi_j^{(k)} \qquad i > k,\; j > k$$

$$b_i^{(k+1)} = b_i^{(k)} \qquad i < k$$

$$b_k^{(k+1)} = b_k^{(k)} - \left(\alpha^{(k)} + a_{kk}^{(k)}\right)\xi^{(k)}$$

$$b_i^{(k+1)} = b_i^{(k)} - a_{ik}^{(k)}\ \xi^{(k)} \qquad i > k$$

Sequential Householder transformation

Assignments: $\alpha^{(k)} = sign\left(a_{kk}\right)\sqrt{a_{kk}^2 \pm \left(a_{(m+1)k}^{(k)}\right)^2}$ $\qquad \beta^{(k)} = \dfrac{1}{\alpha^{(k)} + a_{kk}}$

$$\psi_j^{(k)} = \beta^{(k)}\left(\left(\alpha^{(k)} + a_{kk}\right)a_{kj} \pm a_{(m+1)k}^{(k)}\ a_{(m+1)j}^{(k)}\right) \qquad j > k$$
$$\xi^{(k)} = \beta^{(k)}\left(\left(\alpha^{(k)} + a_{kk}\right)b_k \pm a_{(m+1)k}^{(k)}\ b_{(m+1)}^{(k)}\right)$$

The algorithm:

$$a_{kk}^{'} = -\alpha^{(k)}$$

$$a_{(m+1)k}^{(k+1)} = 0$$

$$a_{kj}^{'} = a_{kj} - \left(\alpha^{(k)} + a_{kk}\right)\psi_j^{(k)} \qquad j > k$$

$$a_{(m+1)j}^{(k+1)} = a_{(m+1)j}^{(k)} - a_{(m+1)k}^{(k)}\psi_j^{(k)} \qquad j > k$$

$$b_k^{'} = b_k - \left(\alpha^{(k)} + a_{kk}\right)\xi^{(k)}$$

$$b_{(m+1)}^{(k+1)} = b_{(m+1)}^{(k)} - a_{(m+1)k}^{(k)}\ \xi^{(k)}$$

where: $a_{(m+1)j}^{(1)} = a_{(m+1)j} \qquad \forall\ j$

$b_{(m+1)}^{(1)} = b_{(m+1)}$

Givens algorithm

Initialization: $A^{(1)} = A$ $\qquad b^{(1)} = b$

The algorithm:

$$a_{ii}^{(k,i+1)} = \sqrt{\left(a_{ii}^{(k,i)}\right)^2 \pm \left(a_{ki}^{(k,i)}\right)^2} \qquad a_{ki}^{(k,i+1)} = 0$$

$$a_{ij}^{(k,i+1)} = \frac{a_{ii}^{(k,i)}a_{ij}^{(k,i)} \pm a_{ki}^{(k,i)}a_{kj}^{(k,i)}}{a_{ii}^{(k,i+1)}} \qquad j > i$$

$$b_i^{(k,\,i+1)} = \frac{a_{ii}^{(k,\,i)} b_i^{(k,\,i)} \pm a_{ki}^{(k,\,i)} b_k^{(k,\,i)}}{a_{ii}^{(k,\,i+1)}}$$

$$a_{kj}^{(k,\,i+1)} = \frac{a_{kj}^{(k,\,i)} a_{ii}^{(k,\,i)} - a_{ki}^{(k,\,i)} a_{ij}^{(k,\,i)}}{a_{ii}^{(k,\,i+1)}} \qquad j > i$$

$$b_k^{(k,\,i+1)} = \frac{b_k^{(k,\,i)} a_{ii}^{(k,\,i)} \pm a_{ki}^{(k,\,i)} b_i^{(k,\,i)}}{a_{ii}^{(k,\,i+1)}}$$

being: $a_{kj}^{(k,1)} = a_{kj}$ $\qquad \forall\ j$

$b_k^{(k,1)} = b_k$

Sequential Cholesky factorization

One observation/equation/row in/out:

$$t_{ii}^{'} = \sqrt{t_{ii}^2 \pm \left(w_i^{(i)}\right)^2} \qquad w_i^{(i+1)} = 0$$

$$t_{ij}^{'} = \frac{t_{ii} t_{ij} \pm w_i^{(i)} w_j^{(i)}}{t_{ii}^{'}} \qquad j > i$$

$$w_j^{(i+1)} = \frac{w_j^{(i)} t_{ii} - w_i^{(i)} t_{ij}}{t_{ii}^{'}} \qquad j > i$$

where: $w_j^{(1)} = a_{(m+1)j}$ $\ \forall\ j$

One parameter/unknown/column in/out

$$t'_{ij} = t_{ij} \qquad \forall i < h \quad \forall j \neq h$$

$$t_{ih} = \frac{c_{ih} - \sum_{k=1}^{i-1} t_{kh} t_{ki}}{t_{ii}} \qquad t_{hh} = \sqrt{c_{hh} - \sum_{k=1}^{h-1} t_{kh}^2} \qquad t_{hj} = \frac{c_{hj} - \sum_{k=1}^{h-1} t_{kh} t_{kj}}{t_{hh}}$$

$$t_{ii}^{'} = \sqrt{t_{ii}^2 - w_i^{(i-h)^2}} \qquad w_i^{(i-h+1)} = 0$$

$$t_{ij}^{'} = \frac{t_{ij} t_{ii} - w_i^{(i-h)} w_j^{(i-h)}}{t_{ii}^{'}} \qquad j > i$$

$$w_j^{(i-h+1)} = \frac{w_j^{(i-h)} t_{ii} - w_i^{(i-h)} t_{ij}}{t_{ii}^{'}}$$

where: $w_j^{(1)} = t_{hj}$ $\qquad j > h$

Matrix inversion lemma

Note: $+ =$ in $\qquad - =$ out

$$\left(Q \pm RST\right)^{-1} = Q^{-1} \mp Q^{-1}R\left(S^{-1} \pm TQ^{-1}R\right)^{-1}TQ^{-1}$$

$$\left(C \pm a^T pa\right)^{-1} = C^{-1} \mp C^{-1}a^T\left(p^{-1} \pm aC^{-1}a^T\right)^{-1}aC^{-1}$$

where: $T^T Te = a^T$

Block Gauss inversion:

"in" "out"

$$\gamma = C^{-1} + \frac{C^{-1}rr^TC^{-1}}{s} + \frac{C^{-1}rr^TC^{-1}}{s\left(s - r^TC^{-1}r\right)} \qquad C^{-1} = \gamma - \frac{ee^T}{s} - \frac{efe^T}{s(s-f)} =$$

$$\rho = -\frac{e}{s} - \frac{ef}{s(s-f)}$$

$$\sigma = \frac{1 - r^T\rho}{s}$$

GRAPH THEORY

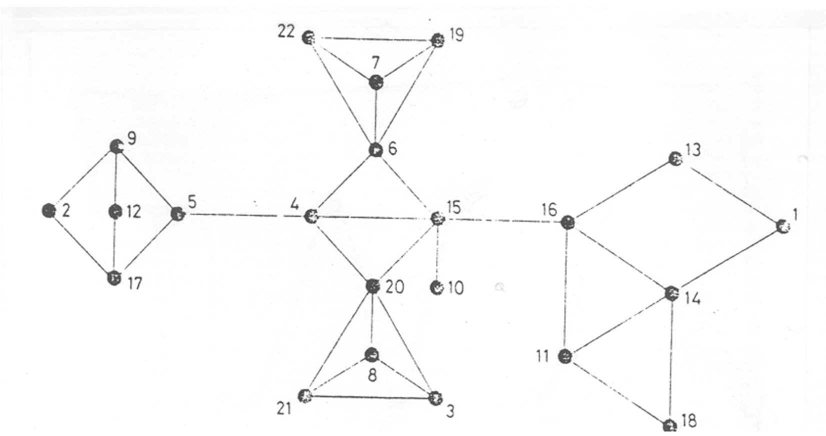
Essential properties:

- A Graph is composed by nodes and arcs, each of these connect two nodes.
- Starting from every node, a tree (or a level structure, which is equivalent) can be built on a graph.
- A tree contains the minimum path (called distance), which connects two given nodes.
- The maximum distance, in a graph, is called diameter.
- The diameter shows the depth of a graph (i.e. the number of levels of its longest level structure).
- The dimension of the related levels is called width of the corresponding level.
- A graph has a width, given by the maximum width of the levels of the best level structure.

Reordering:

- There exists a correspondence between a graph and a square symmetric matrix, whose main diagonal elements represent the nodes of the graph and the non-zero off-diagonal ones the corresponding arcs.
- The best numeration of the nodes of the graph produces a small profile in the corresponding matrix.
- In the planar graph, a small profile is often derived from a small bandwidth.
- A small bandwidth of the matrix is strictly linked to the minimum width in the corresponding graph.
- The minimum width is often derived by the maximum depth (i.e. by the identification of the diameter of the given graph).

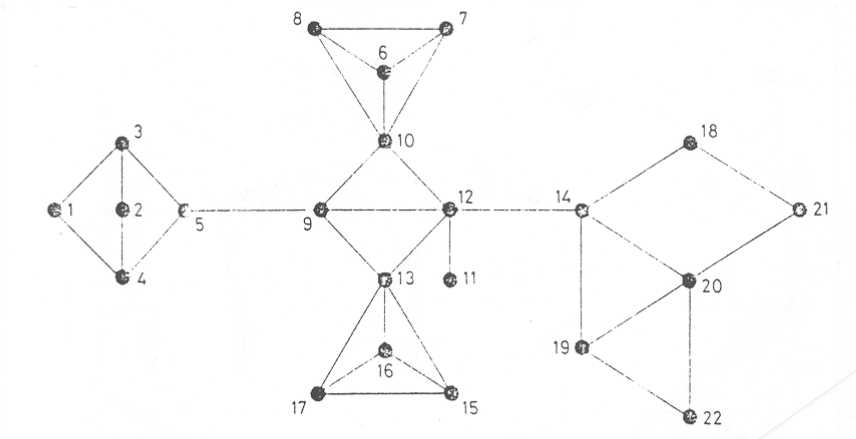
An example: (before the reordering)



Technicalities:

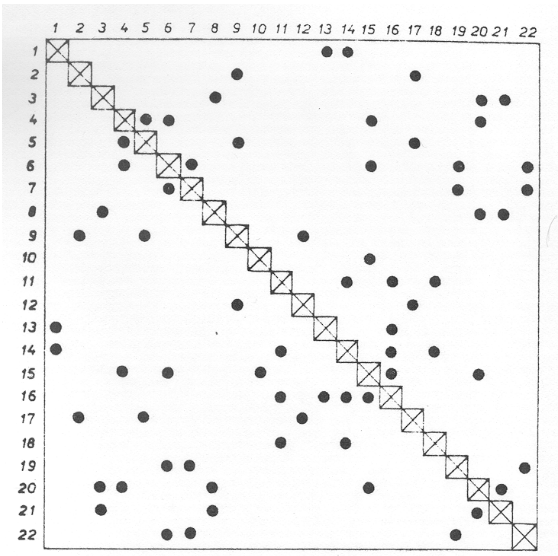
- The diameter is found by iterative search of the longest level structure (whose first element is called root and last elements are called leaves).
- There always exist two level structures, running along the diameter forward and backward.
- It's possible two merge these two level structures in a unique generalized one, which again minimizes its width (the directed strategy puts the nodes, which stay in different levels, with respect to the two original level structures, where the destination levels are more empty).
- The successive numeration proceeds level by level, assigning the number one to a root.
- Inside a level, the numeration proceeds according to the minimum grade (i.e. the number of the locally connected arcs) of the nodes.
- Between two successive levels, the numeration proceeds "parallel" to the previous level.
- If the reverse profile is smaller than the direct one, the numeration is reversed.

An example: (after the reordering)

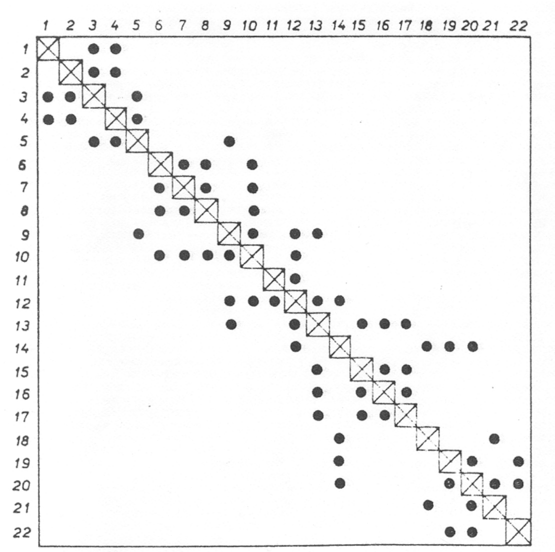


Data summary and the two matrices (before and after the reordering):

	Before	**After**
Matrix Profile	**168**	**79**
Matrix Bandwidth	**18**	**6**



⇒



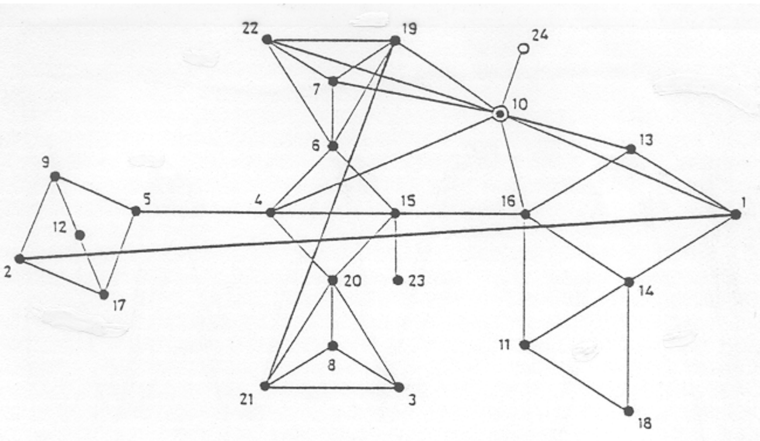
Dissection and reordering

Note: Non-planar graphs show reordering problems, due to nodes with a high grade (i.e. many locally connected arcs) or to long arcs (i.e. arcs linking two nodes, in two different parts of the graph).

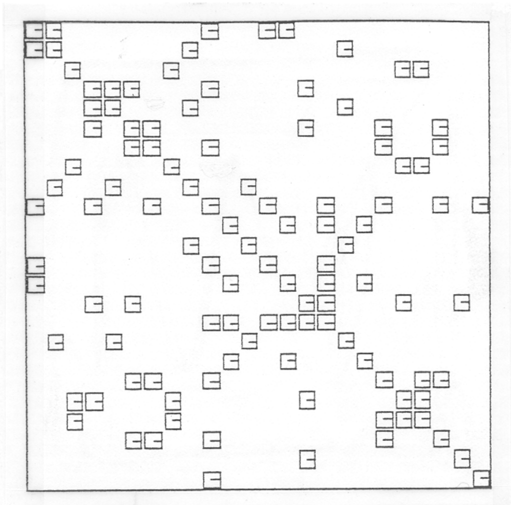
Technicalities:

- Reordering of the whole graph.
- Dissection of the graph, removing high grade node and the long arcs.
- Reordering of the dissected graph only.
- Addition of the right margin with the removed nodes (duplicating the extremes of the long arcs, so that the graph certainly maintains its connection).
- Comparison of the two results (i.e. the profile of the reordering of the whole graph and the profile of the reordering of the dissected graph plus the right margin, with the contribution of the removed nodes) and selection of the smaller one (noting that the bandwidth is only significant in small examples).

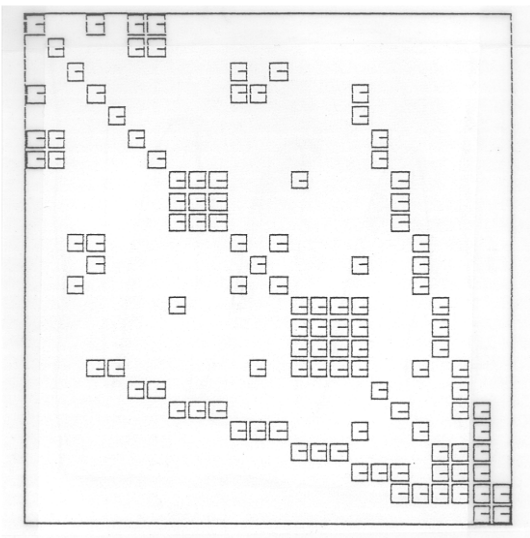
An example:



(before the reordering)

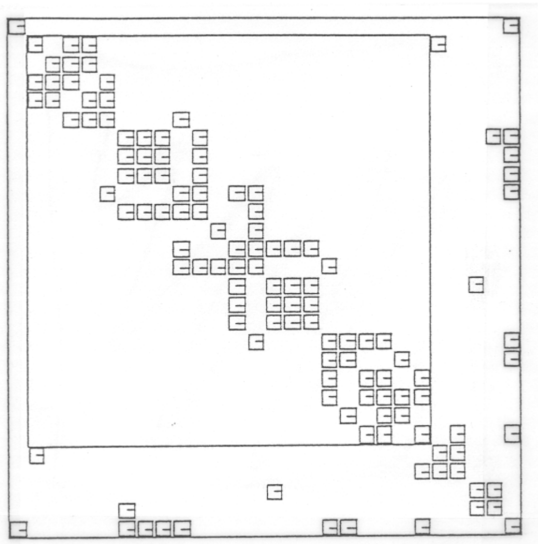


(after the reordering of the whole graph)



⇒

(after the reordering of the dissected graph plus the right margin)



Data summary:

	Before	After (with reordering)	After (with reordering and dissection)
Matrix Profile	202	138	160
Matrix Bandwidth	**19**	**14**	**5 + 5 (right margin)**

PART VI – EXERCISES

TWO EXAMPLES OF DESCRIPTIVE STATISTICS

ONE-DIMENSION STATISTICAL VARIABLE

INPUT DATA

OBSERVATIONS

-0.42 1.13 0.09 -2.01 0.77 0.00 -0.44 0.48 1.91 -0.18 -1.19 -0.24 0.56 0.98 -1.31 0.13

NUMBER OF OBSERVATIONS = 16   NUMBER OF CLASSES = 4

PROCESSING AND RESULTS

MEAN VALUE	STANDARD DEVIATION	SKEWNESS	KURTOSIS
0.02	0.96	-0.2	2.81

INTERVAL = 0.98

HISTOGRAM AND PROBABILITY DENSITY FUNCTION
HISTOGRAM AND CUMULATIVE. DISTRIBUTION FUNCTION

CLASS CENTERS CLASS STANDARD CENTERS ABSOLUTE FREQUENCIES RELATIVE FREQUENCIES (SIMPLE) NORMAL PROBABILITIES	CLASS EXTREMES CLASS STANDARD EXTREMES CUMULATIVE ABSOLUTE FREQUENCIES CUMULATIVE RELATIVE FREQUENCIES CUMULATIVE NORMAL PROBABILITIES
-1.520 -0.540 0.440 1.420	-2.010 -1.030 -0.050 0.930 1.910
-1.604 -0.581 0.442 1.466	-2.116 -1.092 -0.069 0.954 1.977
3 4 6 3	0 3 7 13 16
0.188 0.250 0.375 0.188	0.000 0.188 0.438 0.813 1.000
0.120 0.335 0.358 0.146	0.017 0.137 0.472 0.830 0.976

MEDIAN	MEAN ABSOLUTE VALUE	MEDIAN ABSOLUTE VALUE
0.09	0.74	0.53

TWO-DIMENSION STATISTICAL VARIABLE

INPUT DATA

X / Y	1	2	3	4	MARGINAL FREQUENCIES (Y)	MEAN V. (Y|X)	STANDARD DEV. (Y|X)
1			9		9	3.0	0.0
2	3		2		5	1.8	1.0
3			4	4	8	3.5	0.5
4	1	1		7	9	3.4	1.1
MARGINAL FREQUENCIES (X)	4	1	15	11	31		
MEAN V. (X|Y)	2.5	4.0	1.7	3.6			
STANDARD DEV. (X|Y)	0.9	0.0	0.9	0.5			

PROCESSING AND RESULTS

MEAN VALUE(X) = 2.55 STANDARD DEV.(X) = 1.19
MEAN VALUE(Y) = 3.06 STANDARD DEV.(Y) = 0.95

PEARSON’S INDICES:

$ETA^2(X|Y) = 0.61$ $ETA^2(Y|X) = 0.39$ $ETA^2 = 0.53$

LINEAR REGRESSION:

A(Y(X)) = 0.25 A(X(Y)) = 0.39
B(Y(X)) = 2.43 B(X(Y)) = 1.35 R(X, Y) = 0.31

JOINT NORMAL DISTRIBUTION IN THE INDEPENDET CASE: P(X)×(Y) = 0.85

BEING: XINF = 0.5 XSUP = 4.5 YINF = 0.5 YSUP = 4.5
ZXINF= -1.7 ZXSUP= 1.6 ZYINF= -2.7 ZYSUP= 1.5

BONFERRONI’S INDICES:

B(X) = 0.56 B(Y) = 0.66 B(-1) = 0.61 B(0) = 0.61

Note: Under normal hypothesis, the elementary statistical treatment of the data is limited to one and two dimension analysis, because the only characteristic moments of normal random variables are means and variances, and the only allowed dependence is the linear one, expressed by linear correlation coefficients.

EXAMPLES OF NORMAL
AND DISTRIBUTION-FREE TESTS

INPUT DATA

TWO-DIMENSION STATISTICAL VARIABLE: X, Y

-1.66   0.52 1.85 -0.18 -0.20 1.16   0.50 -0.88 -1.64
-1.63 -1.93 -2.44   1.53 0.64 -1.04 -2.93 -0.19 -0.17

NUMBER OF SAMPLES = 9   NUMBER OF CLASSES = 5

STATISTICS SAMPLING:

MEAN VALUES (X) ; (Y)		STANDARD DEV.'S (X) ; (Y)		CORR. COEF. (X,Y)
-0.0589	-0.9067	1.2020	1.4760	-0.3700

SKEWNESS (X) ; (Y)		KURTOSIS (X) ; (Y)	
0.0333	0.1935	1.5445	1.5278

F(X)	P(X)	F(Y)	P(Y)	F(X,Y)	1	2	3	4	5
0.22	0.23	0.22	0.22	1		0.11		0.11	
0.11	0.21	0.22	0.21	2				0.11	
0.22	0.23	0.11	0.24	3					0.22
0.22	0.18	0.22	0.18	4	0.11	0.11			
0.22	0.15	0.22	0.15	5	0.11		0.11		

CF(X)	CF(Y)	P_CF (X,Y)	1	2	3	4	5
0.22	0.22	1	0.05	0.10	0.12	0.17	0.22
0.33	0.44	2	0.07	0.14	0.18	0.25	0.33
0.55	0.55	3	0.12	0.24	0.30	0.42	0.55
0.77	0.77	4	0.17	0.34	0.42	0.59	0.77
1.00	1.00	5	0.22	0.44	0.55	0.77	1.00

CP(X)	CP (Y)	C(X,Y)	1	2	3	4	5
0.23	0.22	1	-0.05	0.06	-0.02	0.06	-0.05
0.44	0.43	2	-0.02	-0.02	-0.01	0.08	-0.02
0.67	0.67	3	-0.05	-0.05	-0.02	-0.05	0.17
0.85	0.85	4	0.06	0.06	-0.02	-0.05	-0.05
1.00	1.00	5	0.06	-0.05	0.08	-0.05	-0.05

LEGENDA:

F = FREQUENCIES
P_CF = PRODUCT OF CUMULATIVE FREQUENCIES
P = NORMAL PROBABILITIES
CP = CUMULATIVE NORMAL PROBABILITIES
C = CONTINGENCIES

PROCESSING AND RESULTS

PARAMETRIC TESTS ACCORDING TO INDEPENDENCE AND NORMALITY HYPOTHESES

1) GOODNESS OF FIT TEST (BY CHI SQUARE –BASED ON THE SAMPLE VALUES OF MEAN VALUE AND VARIANCE)

SIGNIFICANCE LEVEL: ALFA = 0.02

VAL-EXP(X)	VAL-EXP(Y)	DF	CHI^2-TH(INF) AND	(SUP)
0.83	1.01	2	0.02	9.23

NULL HYPOTHESES: H (0): 0.02

2) GOODNESS OF FIT TEST (BY KOLMOGOROV)
SIGNIFICANCE LEVEL: ALPHA = 0.02

VAL-EXP(X)	VAL-EXP(Y)	DF	KS-TH(SUP)
0.12	0.12	9	0.51

NULL HYPOTHESIS: H (0): Valid

3) NORMALITY TEST(BY PEARSON ET AL.)
SIGNIFICANCE LEVEL: ALPHA = 0.02

VAL-EXP(X)	VAL-EXP(Y)	DF	CHI^2-TH(INF) AND	(SUP)
0.80	0.87	2	0.02	9.23

NULL HYPOTHESES: H (0): 0.02

4) CORRELATION TEST (IN THE NO-CORRELATION CASE)
SIGNIFICANCE LEVEL: ALPHA = 0.02

VAL-EXP	DF	T-TH
1.05	7	3.00

NULL HYPOTHESES: H (0): VALID

5) COMPARISON OF TWO VARIANCES TEST (BY FISHER)
SIGNIFICANCE LEVEL: ALPHA = 0.02

VAL-EXP	DF	F-TH(INF) AND (SUP)	
1.51	8	0.17	6.03

NULL HYPOTHESES: H (0): VALID

6) COMPARISON OF TWO MEAN VALUES TEST (BY T OF STUDENT)
A LEVEL OF SIGNIFICANCE: ALPHA = 0.02

VAL-E	DF	T-TH
1.34	16	2.58

NULL HYPOTHESES: H (0): VALID

7) COMPARISON OF TWO MEAN VALUES TEST (BY WELCH)
SIGNIFICANCE LEVEL: ALPHA = 0.02

VAL-EXP	DF	T-TH
1.34	17	2.57

NULL HYPOTHESES: H (0): VALID

8) VARIANCE TEST (TH. STANDARD DEV. XUY = 1)
SIGNIFICANCE LEVEL: ALPHA = 0.02 (EXP. STANDARD DEV. XUY = 1:38)

VAL-EXP	DF	CHI ^{2} TH(INF) AND (SUP)	
32.22	17	6.40	33.44

NULL HYPOTHESES: H (0): VALID

9) MEAN VALUE TEST (TH. MEAN VALUE XUY) = 0)
SIGNIFICANCE LEVEL: ALPHA = 0.02 (EXP. MEAN VALUE XUY = -0.48)

VAL-EXP	DF	T-TH
1.49	17	2.57

NULL HYPOTHESES: H (0): VALID

10) POWER CURVE OF THE MEAN VALUE TEST
SIGNIFICANCE LEVEL: ALFA = 0.02

H (0): M (XUY) = 0

H (1): M (XUY) = K $\times$ DELTA(M (XUY)) WITH DELTA (M (XUY)) = 1.22/SQRT(2N)

0	$1 \times$ DELTA (M (XUY))	$2 \times$ DELTA (M (XUY))	$3 \times$ DELTA (M (XUY))
0.00	0.89	1.77	2.66 STANDARD ABSCISSA
0.02	0.08	0.29	0.63 POWER (1 – BETA)

DISTRIBUTION-FREE TESTS

1) INDEPENDENCE TEST (BY CHI SQUARE)
SIGNIFICANCE LEVEL: ALFA = 0.02

VAL-EXP	DF	CHI^2-TH(INF) AND (SUP)	
20.25	16	5.80	32.03

NULL HYPOTHESES: H (0): VALID

2) INDEPENDENCE TEST (BY KOLMOGOROV)
SIGNIFICANCE LEVEL: ALFA = 0.02

VAL-EXP	DF	KS-TH(SUP)
0.19	9	0.51

NULL HYPOTHESES: H (0): VALID

3) SPEARMAN CORRELATION TEST (IN THE NO-CORRELATION CASE)
SIGNIFICANCE LEVEL: ALPHA = 0.02 (R = -0.45)

VAL-EXP	DF	T-TH
1.33	7	3.00

NULL HYPOTHESES: H (0): VALID

4) MANN – WHITNEY RANK TEST (COMPARISON OF CENTRAL VALUES)
SIGNIFICANCE LEVEL: ALPHA = 0.02

SUM(RANK (X))	VAL-EXP	Z-TH
98.00	1.10	2.33

NULL HYPOTHESES: H (0): VALID

5) SIEGEL – TUKEY RANK TEST (COMPARISON OF DISPERSION)
SIGNIFICANCE LEVEL: ALPHA = 0.02

SUM(RANK (X))	VAL-EXP	Z-TH
77.00	-0.75	2.33

NULL HYPOTHESES: H (0): VALID

6) THOMPSON SIGN TEST (COMPARISON OF CENTRAL VALUES)
SIGNIFICANCE LEVEL: ALPHA = 0.02

F(+ DIF)	VAL-EXP	Z-TH
0.55	0.30	2.33

NULL HYPOTHESES: H (0): VALID

7) THOMPSON SIGN TEST (COMPARISON OF DISPERSION)
SIGNIFICANCE LEVEL: ALPHA = 0.02

F(+ DIF)	VAL-EXP	Z-TH
0.55	0.30	2.33

NULL HYPOTHESES: H (0): VALID

TWO EXAMPLES OF LEAST SQUARES PROBLEMS

LINEAR INTERPOLATION IN THE SPACE DOMAIN

S = A + BX + CY + DZ + H BEING: H = – 0.60, A GIVEN CONSTANT

INPUT DATA

IND. VAR. X	IND. VAR. Y	IND. VAR. Z	OBSERVATIONS S
0.71	0.62	0.17	-1.50
-0.71	0.62	0.17	0.11
0.71	-0.62	0.17	0.57
0.71	0.62	-0.17	-0.35
-0.71	-0.62	0.17	0.26
-0.71	0.62	-0.17	-0.06
0.71	-0.62	-0.17	-0.55
-0.71	-0.62	-0.17	-0.54

NUMBER OF OBSERVATIONS AND EQUATIONS = 8
NUMBER OF PARAMETERS AND UNKNOWNS = 4

PROCESSING AND RESULTS

Note: The following results imply elementary operations of linear algebra, like:

- algebraic sum of matrices;
- product of matrices and product of a matrix times a scalar;
- computation of trace and determinant of a square matrix;
- matrix inversion.

DESIGN MATRIX A				KNOWN VECTOR B
1.00	0.71	0.62	0.17	0.90
1.00	-0.71	0.62	0.17	-0.71
1.00	0.71	-0.62	0.17	-1.17
1.00	0.71	0.62	-0.17	-0.25
1.00	-0.71	-0.62	0.17	-0.86
1.00	-0.71	0.62	-0.17	-0.54
1.00	0.71	-0.62	-0.17	-0.05
1.00	-0.71	-0.62	-0.17	-0.06

NORMAL MATRIX				NORMAL KNOWN VECTOR D
8.00	0.00	0.00	0.00	-2.74
	4.03	0.00	0.00	1.14
SYM.		3.08	0.00	0.95
MATRIX			0.23	-0.16

SOLUTION X	STANDARD DEVIATION (SD) OF THE SOLUTION
0.34	0.26
-0.28	0.36
-0.31	0.41
0.69	1.51

RESIDUALS V	SD OF THE RESIDUALS	ESTIMATES S	SD OF THE ESTIMATES
0.97	0.51	-0.53	0.51
-0.24	0.51	-0.13	0.51
-0.72	0.51	-0.15	0.51
-0.42	0.51	-0.77	0.51
-0.01	0.51	0.25	0.51
-0.31	0.51	-0.37	0.51
0.17	0.51	-0.38	0.51
0.56	0.51	0.02	0.51

SIGMA ZERO 0.73 CONDITION NUMEBER 0.03

LOCAL REDUNDANCIES

0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50

CUBIC POLYNOMIAL INTERPOLATION IN THE TIME DOMAIN

$S = A + BT + CT^2 + DT^3 + H$ BEING: H = 1.46, A GIVEN CONSTANT

INPUT DATA

IND. VAR. T	OBSERVATIONS S
0.17	-1.50
-0.17	0.11
0.34	0.57
-0.34	-0.35
0.51	0.26
-0.51	-0.06
0.68	-0.55
-0.68	-0.54

NUMBER OF OBSERVATIONS AND EQUATIONS = 8
NUMBER OF PARAMETERS AND UNKNOWNS = 4

PROCESSING AND RESULTS

DESIGN MATRIX A				KNOWN VECTOR B
1.00	0.03	0.17	0.00	2.96
1.00	0.03	-0.17	-0.00	1.35
1.00	0.12	0.34	0.04	0.89
1.00	0.12	-0.34	-0.04	1.81
1.00	0.26	0.51	0.13	1.20
1.00	0.26	-0.51	-0.13	1.52
1.00	0.46	0.68	0.31	2.01
1.00	0.46	-0.68	-0.31	2.00

NORMAL MATRIX				NORMAL KNOWN VECTOR D
8.00	1.73	0.00	0.00	13.74
	0.59	0.00	0.00	3.00
SYM.		1.73	0.59	-0.20
MATRIX			0.24	-0.07

SOLUTION X	STANDARD DEVIATION (SD) OF THE SOLUTION
-1.70	0.49
-0.09	1.81
0.10	1.67
0.03	4.53

Note: A simple reordering has been done, writing the coefficient of t^2 before the coefficient of t, so that the normal matrix becomes of block diagonal form.

RESIDUALS V	SD OF THE RESIDUALS	ESTIMATES S	SD OF THE ESTIMATES
1.28	0.66	-0.22	0.52
-0.37	0.66	-0.26	0.52
-0.78	0.65	-0.21	0.54
0.07	0.65	-0.28	0.54
-0.47	0.69	-0.21	0.48
-0.26	0.69	-0.32	0.48
0.35	0.30	-0.20	0.78
0.18	0.30	-0.36	0.78

SIGMA ZERO 0.84 CONDITION NUMEBER 0.00

LOCAL REDUNDANCIES

0.61 0.61 0.59 0.59 0.67 0.67 0.13 0.13

THREE EXAMPLES OF GEODETIC LEAST SQUARES PROBLEMS

ADJUSTMENT OF TWO LATTICE STRUCTURES

A) FUNCTIONAL MODEL: FINITE DIFFERENCES OF FIRST ORDER

$$D(I,J)=ALPHA\times Z(J)-BETA\times Z(I)+GAMMA$$

ALPHA = 5 BETA = 2 AND GAMMA = 4, THREE GIVEN CONSTANTS

STOCHASTIC MODEL:

INDEPENDENT OBSERVATIONS OF EQUAL VARIANCES
WEIGHT OF THE CONSTRAINT (OF THE UNKNOWN 1) EQUAL TO 10000

INPUT DATA

I	J	D(I, J)
1	2	1.81
2	3	4.36
3	4	3.61
4	1	4.62
1	3	3.23
2	4	5.15

NUMBER OF OBSERVATIONS, PSEUDO-OB.'S AND EQUATIONS = 7
NUMBER OF PARAMETERS AND UNKNOWNS = 4

Note: This problem and the following one present a network structure, where the rows of the design matrix contain the observations of the network, which are like the arcs of a graph and the columns of the design matrix contain the unknowns, which are like the knots of the same graph.

PROCESSING

DESIGN MATRIX A				KNOWN VECTOR B	WEIGHTS
-2	5			2.19	1
	-2	5		-0.36	1
		-2	5	0.39	1
5			-2	-0.62	1
-2		5		0.77	1
	-2		5	-1.15	1
1				0	10.000

B) FUNCTIONAL MODEL: FINITE DIFFERENCES OF SECOND ORDER

$$D(I,J,K)=ALPHA\times Z(I) - 2\times BETA\times Z(J)+GAMMA\times Z(K)+DELTA$$

ALPHA = 10 BETA = 4 GAMMA = 2 AND DELTA = 5, FOUR GIVEN CONSTANTS

STOCHASTIC MODEL:

INDEPENDENT OBSERVATIONS OF EQUAL VARIANCES
WEIGHT OF THE CONSTRAINTS (OF UNKNOWNS 1 AND 2) EQUAL TO 10000

INPUT DATA

I	J	K	D(I, J, K)
3	1	2	4.32
1	2	4	6.91
4	3	1	4.67
2	4	3	4.26
5	3	4	5.36
3	4	6	7.79
6	5	3	3.77
4	6	5	4.93

NUMBER OF OBSERVATIONS, PSEUDO-OB.'S AND EQUATIONS = 9
NUMBER OF PARAMETERS AND UNKNOWNS = 6

PROCESSING

DESIGN MATRIX A						KNOWN VECTOR B	WEIGHTS
-8	2	10				0.68	1
10	-8		2			-1.91	1
2		-8	10			0.33	1
	10	2	-8			0.74	1
		-8	2	10		-0.36	1
		10	-8		2	-2.79	1
		2		-8	10	1.23	1
			10	2	-8	0.07	1
1						0	10.000
	1					0	10.000

<u>Note:</u> The problems with a network structure present always a rank defect, whose size depends from their proper characteristic (e.g. the rank defect is equal to one in the finite difference problems of first order and equal to two in the finite difference problems of second order). Notice that the rank defect is solved by using an equal number of suitable constrains, which can be substituted by pseudo-observations, with a very high weight.

<u>SURFACE RECONSTRUCTION</u>

FUNCTIONAL MODEL: FINITE ELEMENTS INTERPOLATION

$S = A + B \times X + C \times Y + H$ IF $X^2 + Y^2 < 4$
$S = D + E \times X + F \times Y + K$ IF $X^2 + Y^2 > 4$

$H = -0.62$ AND $K = 2.33$, TWO GIVEN CONSTANTS

INDEPENDENT OBSERVATIONS OF EQUAL VARIANCES
NO ADDITIONAL CONSTRAINTS ARE REQUIRED

INPUT DATA

IND. VAR. X	IND. VAR. Y	OBSERVATIONS S
0.16	0.16	0.57
– 0.16	0.16	0.78
– 0.16	– 0.16	0.92
0.16	– 0.16	– 0.52
2.57	2.57	– 0.70
– 2.57	2.57	– 1.87
– 2.57	– 2.57	– 0.17
2.57	– 2.57	– 0.48

PROCESSING

DESIGN MATRIX A						KNOWN VECTOR B	WEIGHTS
1	0.16	0.16				-1.19	1
1	-0.16	0.16				-1.40	1
1	-0.16	-0.16				-1.54	1
1	0.16	-0.16				-0.10	1
			1	2.57	2.57	3.03	1
			1	-2.57	2.57	4.20	1
			1	-2.57	-2.57	2.50	1
			1	2.57	-2.57	..2.81	1

Note: The product of two matrices and the inversion of matrices are easy standard operations of linear algebra, especially if their dimensions are rather small. Nevertheless if the dimensions are equal to two or three, close formulas can be written, as follows.

Product:

$$AB = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix} = C$$

being: $c_{ij} = \sum_{k=1}^{4} a_{ik} b_{kj} \qquad i = 1,3 \qquad j = 1,2$

Inversion: (2×2 matrix)

$$Q = \begin{bmatrix} a & c \\ sym. & b \end{bmatrix}$$

$$det\,Q = ab - c^2 \qquad\qquad Q^{-1} = \frac{1}{det\,Q} \begin{bmatrix} b & -c \\ sym. & a \end{bmatrix}$$

(3×3 matrix)

$$Q = \begin{bmatrix} a & d & e \\ & b & f \\ sym. & & c \end{bmatrix}$$

$$det\,Q = abc + 2def - af^2 - be^2 - cd^2$$

$$Q^{-1} = \frac{1}{det\,Q} \begin{bmatrix} bc - f^2 & -(cd - ef) & df - be \\ & ac - e^2 & -(af - de) \\ sym. & & ab - d^2 \end{bmatrix}$$

PART VII – DATA PROCESSING MEETS HUMAN SCIENCES

THE ANCIENT WORLD

The human species, originated from anthropoid primates, is widespread. However, neither subspecies nor hybrids have been observed since at least 10.000 years.

- Palaeolithic and Neolithic periods: discovery of fire, plough, wheel, mill, wagon.
- Neolithic revolution: domestication of animals and plants.
- Prehistory: weaving, handcrafted pottery and metal tools.

Signs and symbols:

- Neanderthal;
- Camuni and Celts (Stonehenge);
- Aztecs, Incas, Easter Island, etc.

- Language is a human feature.
- Animals communication exists as well: hymenoptera chemical communication (bees, wasps and ants);

  cetacean's singing (whales and dolphins);

  monkeys' different defence whistles (against leopards, snakes and eagles),

  Nonetheless, carnivorous mammals, equines, elephants, birds of prey, ravens and parrots are equally intelligent animals.

Origin of writing:

- Sumerian Craetulae.
- Assyrian, Babilonyan and Persian tablets.
- Egyptian ideograms.
- Phoenician cuneiform script.
- Cretan linear scripts: A (Minoan: un-translated, till now) and B (Mycenaean: recognized as ancient Greek).
- Ancient languages used in Greece, Anatolia, Magna Graecia and in Rome.
- Other ancient written languages in India, China and in Maya civilization.

  - Other derivate languages exist (such as Romance, Saxons and Slavic languages); moreover Pitney and Creole languages.
  - The translation of seventy interpreters (known as Septuagint) of the Old Testament from Jewish to Koinè Greek is an example of ancient linguistics.

Ancient Mathematics:

- Mesopotamia / fertile half-moon (where Semitic languages are spoken): arithmetic and astronomy.
- Egypt (where Hamitic languages are spoken, as for instance Somali, Ethiopic, Chadic and Berber, even though they are pretty different from one another): geometry and astronomy.
- Greek and Hellenic area: geometry: Thales, Pythagoras and Euclid;
  Eratosthenes of Cyrene, Archimedes, Apollonius of Perga;
  Hero of Alexandria and Pappus of Alexandria.

  Euclid postulates:

  - A straight line segment can be drawn joining any two points
  - Any straight line segment can be extended indefinitely in a straight line.
  - Given any straight line segment, a circle can be drawn having the segment as radius and one endpoint as centre.
  - All right angles are congruent.
  - Given any straight line and a point not on it, there "exists one and only one straight line which passes" through that point and never intersects the first line, no matter how far they are extended (this statement is the modern version of the original fifth Euclid Postulate).

Astronomy:

- Heliocentrismus: Pytagoras ed Aristharcus of Samos (with Archimedes).
- Geocentrismus: (with Aristotle), Timocharis, Hipparchus of Nicaea, Ptolemy of Alexandria.

Travel and cartography:

- Argonauts endeavour and Odysseus’ travels.
- Pillars of Hercules and Atlantis myth.
- Phoenician circumnavigation of Africa and Herodotus tales.
- Alexander the Great conquests and Hipparchus of Nicaea hypothesis.
- Eratosthenes of Cyrene (measure of earth radius) and Ptolemy of Alexandria (map of the Ecumene).

Ancient philosophy:

- Myth poems: Homer and Hesiod.
- Pre-Socratic: Anaximander, Heraclitus, Empedocles and Democritus.
- Sophists and Socrates.
- Plato, the Academy and the World of Ideas.
- Aristotle, the Lyceum (or Peripatetic school), the Physics and the Metaphysics.
- Stoicism (in Rome: latter Virgil, Seneca and Marcus Aurelius).
- Epicureans (in Rome: Lucretius, Horace and earlier Virgil) and the Scepticism (in Rome: Cicero).

- The Stoicism is an opposition philosophy in Greece and the leading class ideology during the Roman empire.
- The Christianity (which starts as an Hebraic heresy and becomes pagan religion, with Paulus of Tarsus) gets myths and traditions from Egypt and from other pagan areas; it's the heir of Stoicism with openings towards lower classes, but less tolerance.

Ancient Medicine:

- Chiron the Centaur and Asclepius.
- Hippocrates' science and Galen's techniques.
- Medicine of the Peripatetic school in Alexandria.

THE MIDDLE AGES AND MODERN AGE

From the fall of the ancient world to modern age:

- Republics and empires between Greece and Rome.
- Barbaric invasion and/or ethnic migration.
- Augustine of Hippoa, the Patristics and the Gnosticism (derived from Persian Dualism).
- City abandonment and construction of walls around villas and churches, forming castles and abbeys (with Romanesque and Byzantine styles).
- Monasticism and feudalism.
- Arabs, Persians and Islam of Turks (Sunni and Shia Islam).
- "Alloderi", merchants and city rebirth (and of the Maritime Republics).
- Black death and the "Signorie di Banno".
- Principality and sovereign nations (Italian case, the birth of Switzerland and Nederland).
- Thomas Aquinas and the Byzantine, Persian and Arabic heritage in the Arabic Spain (with Avicenna, Averroes, Maimonides and Ramon Llull).
- Aristotle's Thomism.
- Renaissance Neo-Platonism (with mathematical and technological innovation, printing among others).
- Italian Renaissance artistic development (between Gothic and Baroque).
- Heretical movements, Mendicant orders, Lutheranism, the Counter-reformation and Religious wars
- The new science.
- Age of Enlightenment and Age of Reformations (British case and the birth of the United State of America).
- French Revolution, Napoleon and the European Restoration.

    - Italian and other European countries (e.g. Greece, Belgium, Hungary, Poland, Eire and Balkan countries) independence and/or unification
    - Colonialism and Imperialism.
    - First World War.

- Nationalism and Totalitarianism.
- World War Two.
- The Cold War, Neo-Imperialism and European Communitarians.
- Globalizations and present crisis (environmental, political and economical).

Travels:

- Marco Polo, Ibn Battuta (Arab) and Zheng He (Chinese).
- Vivaldi brothers and the Portuguese sailors.
- Columbus (followed by Cortes and Pizarro).
- Bartolomeo Dias and Vasco da Gama.
- Vespucci and Magellan (as well as Caboto and Verrazzano).
- Tasman and Cook.
- Nobile, Amundsen and Scott.

Philosophy:

- English Scholasticism (Roger Bacon, Duns Scotus).
- Humanists and the people of the Italian Renaissance (mainly Machiavelli).
- Bacon and the new Science.
- Descartes and the Rationalism.
- British Empiricists (Hobbes, Locke, Berkeley, Hume, Bentham).
- Liberal economists, but not laissez-faire (Smith and Ricardo).
- Pre-Enlightenment and Spinoza.
- French Enlightenment agents (Voltaire, Montesquieu, Rousseau, Diderot and D'Alembert).
- Vico, Herder and Goethe.
- Kant and critic synthesis.

  - The "A priori" of time and space (according to Euclid's geometry and Newton's physics).
  - Categories:

Categories:	quantity:	universal, peculiar and singular; i.e.: unity, plurality and totally;
	quality:	affirmative, negative and infinite i.e.: positivity, negativity and limitation;
	relationship:	categorical, hypothetic and disjunctive, i.e.: substantiality, causality, reciprocity;
	modal:	problematic, supportive and apodictic, i.e.: possibility, reality and need.

  - A priori idea of God, Soul and World (i.e. science at that time without object).
  - Mathematics is analytical and its base is a priori synthetic.
  - Moral answers to categorical imperative (every human being is a goal and not a mere mean), but politics can be freely bargained.

Mathematics and Mathematical physics:

- Fibonacci and Arab digits with the Indian / Arab (and Maya) zero.
- Luca Pacioli and mathematic operations signs.
- Trigonometry (at the Leonardo Da Vinci's machine time).
- Algebraic experts in Bologna and third and fourth grade equations (then Ruffini, Abel and Galois).
- Descartes and analytic geometry.
- Desargues and projective geometry (between the invention of prospective and graphic informatics).
- Fermat, Pascal, the Bernoullis and the game theory.
- Leibniz, Newton and differential computation (while the defined integrals had already been used by Archimedes).

  - Newton deals with optics and mechanics as well:

    Newton's First Law of Motion (inertia, already from Galileo);
    Newton's Second Law of Motion (force and acceleration);
    Newton's Third Law of Motion (action and reaction).

- Euler, mathematical geometry, algebra and mathematical analysis.
- Lagrange, Hamilton and analytical mechanics.
- Legendre, Gauss, differential geometry and statistics.
- Cauchy and mathematical analysis.

Music:

- Byzantine Psalter, Gregorian chants and the Laudi.
- Troubadour and Trouvere, the Ars Nova, polyphony and madrigals
- Bach's Well Tempered Clavier, in addition to Baroque music.
- Mozart's mathematical counterpoint.
- Beethoven's great symphonies.

Sciences:

- Astronomy: Copernicus (with Giordano Bruno), Kepler and Galileo;
  Newton (Universal gravitation), Herschel (with Abbot Boscovich).

- Geodesy e cartography: Mercator;
  the Cassinis, Newton, Clairaut and the shape of the Earth.
  Expeditions in Lapland and Peru, to measure meridian arcs;
  Lambert and Gauss;
  Legendre and spherical trigonometry;
  Stokes and the deviation from vertical;

Molodensky and the physical surface of the Earth.

- Physics: Hooke's elasticity Law;
  Bernoulli's and Stevin's hydraulic.
  Torricelli's and Malebranche's fluid mechanics;
  Galvani's and Volta's first electromagnetism studies.

- Chemistry: Boyle and the Phlogiston theory;
  Lavoisier (and the oxygen invention), Liebig (organic chemistry);
  Mendeleev (elements classification), Mendel (genetics).

- Biology: Linnaeus (natural classification), Spallanzani (sexual reproduction);
  Lamarck, Darwin (evolutionist theory).

- Medicine: Vesalius, Harvey (blood circulation), Malpighi (microscopic analysis).
  Jenner (vaccinations), Semmelweis (hygiene), Pasteur (microbiology).

- Linguistic: Valla (correct document dating: De falso credita et ementita donatione Costantini).
  Luter (Bible translation from Greek, emending Saint Jerom's Vulgata);
  Discovery of Sanskrit;
  Rosetta's stone translation;
  Von Humbuldt and the Indo-European languages.

CONTEMPORARY AGE AND PRESENT WORLD

The various branches of knowledge grow apart and become different, but tend to have overlaps and cross-breed with each other.

Philosophy:

- Idealism: mainly Hegel.
- "Spiritualism": Kierkegaard, Schopenhauer, Nietzsche and Bergson.
- Marxism: a liberal revolutionary: Feuerbach.
  Founding fathers: Marx (who asserts not to be a Marxist) and Engels;
  Schools: Second International: labour and social democratic;
  Third International: communist (with Luxemburg and Lenin);
  degeneration: Trotsky's permanent revolution;
  Stalin's bureaucratic totalitarianism;
  Two contribution: Gramsci's historic block;
  Brandt's peaceful coexistence and distension.
- Historicism: Croce (political fracture with Gentile and cultural closure with Enriques).

- ❑ Neo-Kantism: mainly Cassirer (with Mach logical positivism)
  going alongside, though separately, with: Husserl's phenomenology;
  Gestalt's philosophy.
- ❑ Critic empiricism: mainly Reichenbach and Carnap.
- ❑ Existentialism: mainly Heidegger (a shameless Nazi), Jaspers and Sartre.
- ❑ Structuralism: Foucault and the Frankfurt School with: Adorno, Horkheimer and then Habermas.
- ❑ Pragmatism: (Americans) Peirce, James, Dewey and then Rawls;
  (European "predecessors") Bentham, Tocqueville, Stuart Mill and Weber.
- ❑ Falsifiability: Popper, then Kuhn, Feyerabend, Lakatos
  (after then the American analytical philosophy).

<u>Logics:</u>

- ❑ Boole and Peano;
- ❑ Frege and Russell;
- ❑ Hilbert, Tarski, Gödel and Church.

<u>Mathematics:</u>

- ❑ Non Euclidean Geometries: Saccheri, Legendre and Gauss;
  Bolyai, Lobachevsky and Riemann.
- ❑ Topology: Euler, Jordan, Klein, Poincare, Hausdorf.
- ❑ Arithmetic and algebra: Krönecker, Dedekin and Cantor.
- ❑ Mathematical analysis: Fourier, Laplace and Weierstrass;
  Hilbert, Lebesgue and Bourbakists.
- ❑ Statistics: Bayes, Galton, Pearson, Fisher and Tukey.

<u>Physics:</u>

- ❑ Thermodynamic: Carnot, Joule and Lord Kelvin.
- ❑ Electromagnetism: Faraday, Maxwell, Hertz and Lorenz.
- ❑ Energetics, atomistics, radioactivity: Mach, Boltzmann and the Curies.
- ❑ Relativity Theory: Poincaré and Einstein.
- ❑ Quantum theory: Planck, Rutherford, Bohr, De Broglie, Dirac, Schrödinger,
  Pauli, Heisenberg and the Via Panisperna group.

  - ❑ Quantum gravity, string theory, parallel universes, etc.

<u>Informatics:</u>

- ❑ Abaci

- ❑ Logarithm tables
- ❑ Pascal's calculator
- ❑ The "Brunschwiga" mechanical calculator.
- ❑ Punched cards.
- ❑ Relays for phones.
- ❑ Wiener's cybernetics.
- ❑ Turing's machine.
- ❑ Von Neumann's computer.
- ❑ Transistors.
- ❑ Chip and microchip.
- ❑ Internet.

  - ❑ Artificial Intelligence (a dream).
  - ❑ Knowledge engineering (present reality).

<u>Linguistic:</u>

- ❑ De Saussure (from comparative to structural grammar).
- ❑ Wittgenstein with formal and spoken language.
- ❑ Trubeckoj and phonemes.
- ❑ Bloomfield and syntagma.
- ❑ Chosmky (from structural to generative and transformational grammar).

  - ❑ Innatism: of deep grammar;
of symmetries;
of first natural numbers.

  - ❑ What one says (according to Davidson).
  - ❑ What makes a man and a woman different from computers (according to Searle).

<u>Science:</u>

- ❑ Biology: Watson's and Creek's DNA, and Dulbecco's genome.
- ❑ Medicine: Fleming (penicillin), Sabin (new vaccinations), Barnard (transplants).
- ❑ Psychology: Freud (unconscious), Jung (archetype), Piaget (childhood psychology).

<u>The conquest of space (locomotion and energies):</u>

- ❑ The Montgolfiers.
- ❑ The Wright brothers.

- Piccard and the troposphere.
- The Sputnik and the dark side of the moon.
- The Apollo Project and the moon.
- Venus and Mars.
- The Pioneer.

  - The steamboat
  - The internal combustion engine (Otto).
  - The Diesel engine.
  - Nuclear energy.
  - Renewable energy: solar, wind, geothermic and marine.

<u>Literature art and music:</u>

- Romanticism, Verism, Hermeticism and contemporary literature.
- Impressionism, secession, expressionism and abstract art / minimalism.
- Grand opera (in addition to Symphonic music), café chantant, dodecaphony and modern contaminations.

  - Telegraph and telephone.
  - Cinema, radio and television.
  - Fax and Internet.

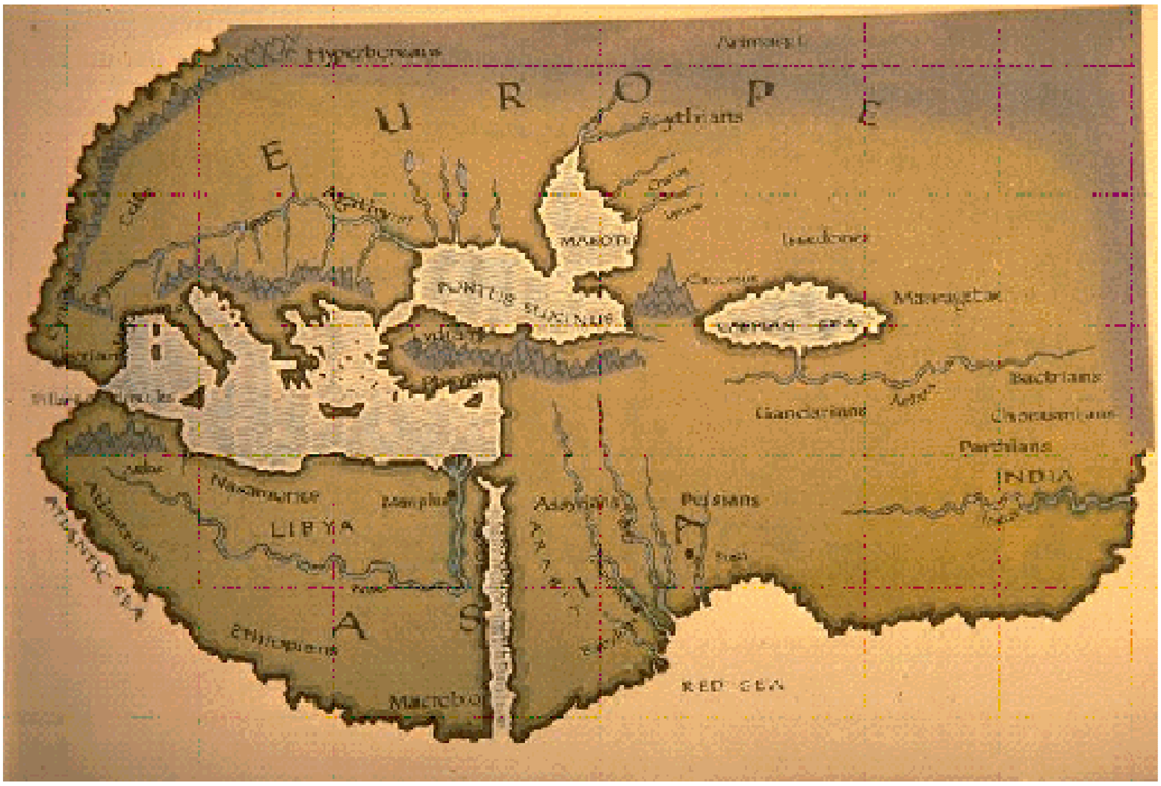
<u>Pedagogy and ethics:</u>

- Knowledge and learning maze (for a confederation and not an empire).
- Out from old morals (even from the protestant ethic of responsibility and firm belief).
- Ethics as an etiquette (the Beauty towards the Good and so, through the Right, in seek of the Truth).
- A freely bargained and shared moral (but not too far away, due to convenience).
- No one is not guilty (and the rights of the second-to-last).
- Secularization (with scepticism and mild relativism) and the (eclectic) valorisation of traditions.
- Being able to be in minority, because pluralism is not chaos, but a great value.

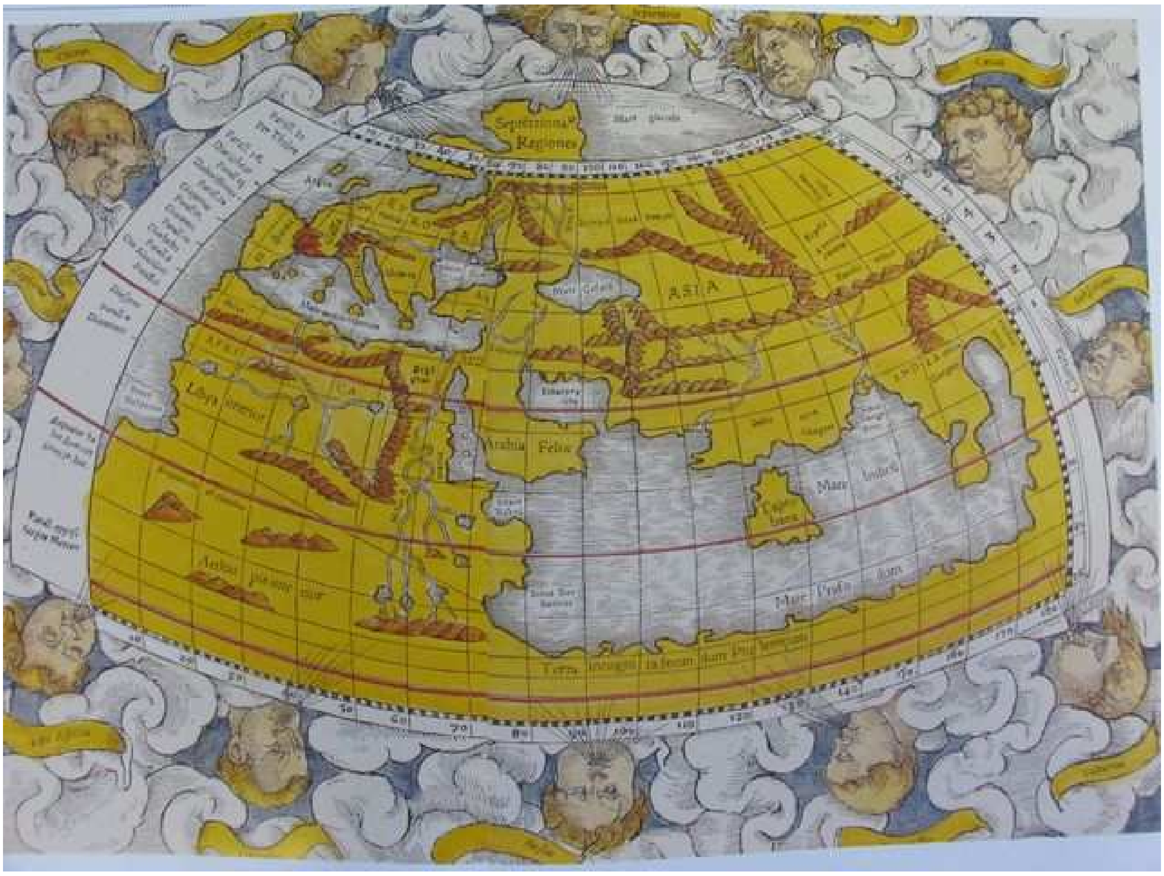
<u>Time and Place of Utopia:</u>

- Gold Age and Eden.
- “The land of milk and honey”.
- The Enchanted forest.
- The ideal city and the utopia threshold.
- The sky freed from monsters and tolerance as a means of measure.
- The futuristic metropolis and the happy island.
- The crushed utopia (of crossbreed and cultural polytheism).

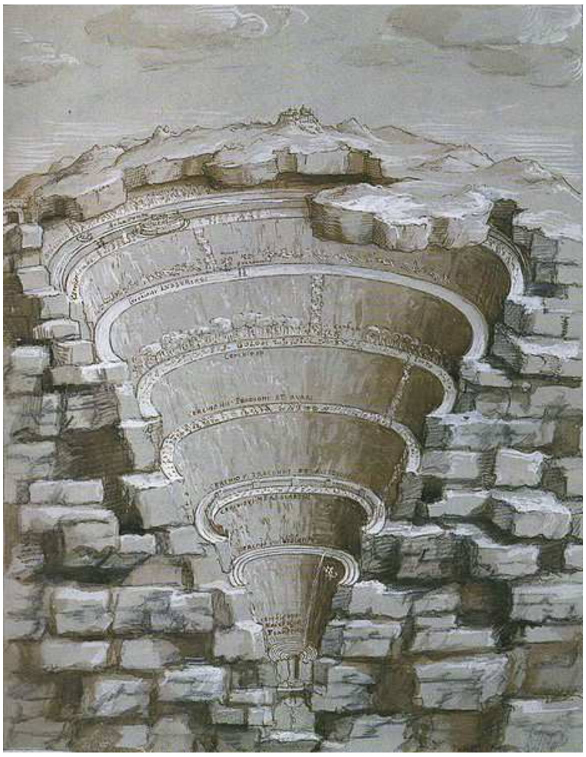
A map collection along the path of history



Herodotus' Ecumene map

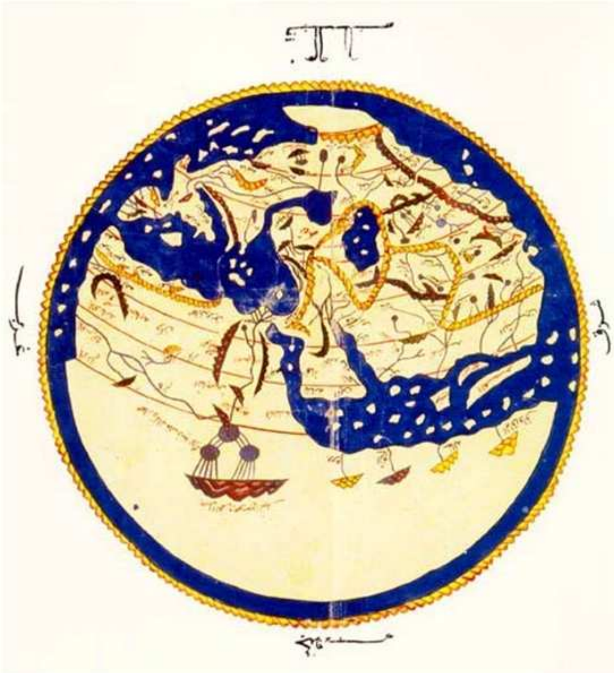


Ptolomy's Ecumene map



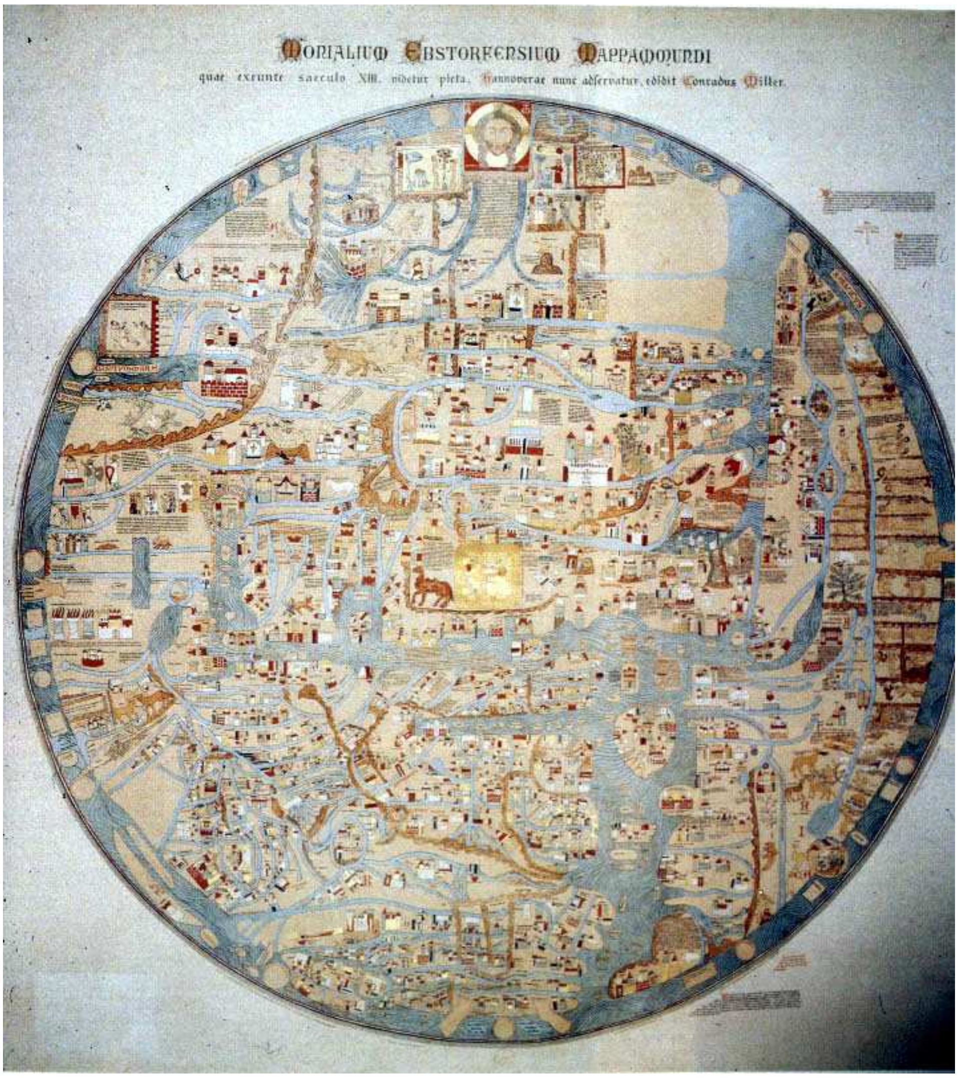
Jan Van der Straet, Dante Alighieri's Hell

Ambrosius Holbein, Thomas More's Utopia

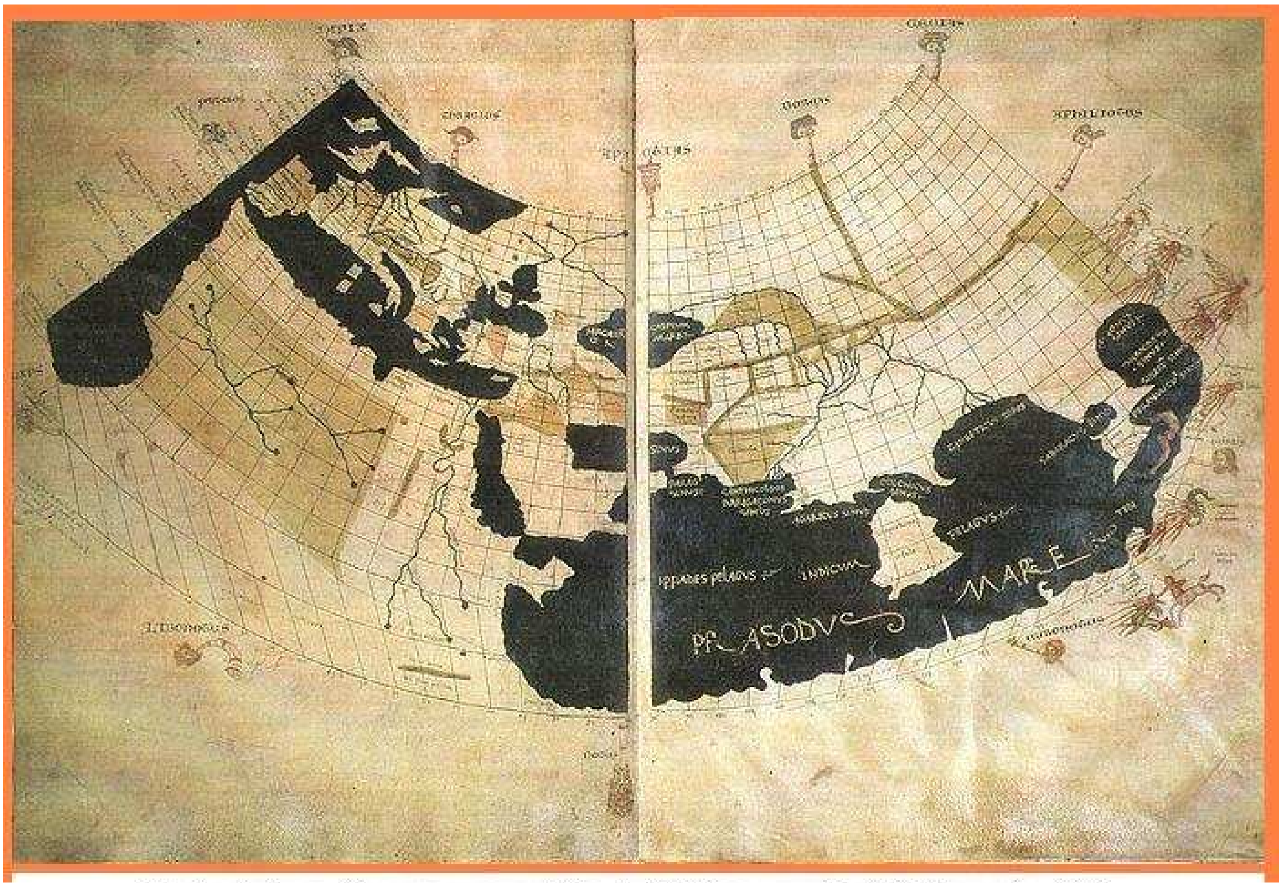


Isidore of Siville's map

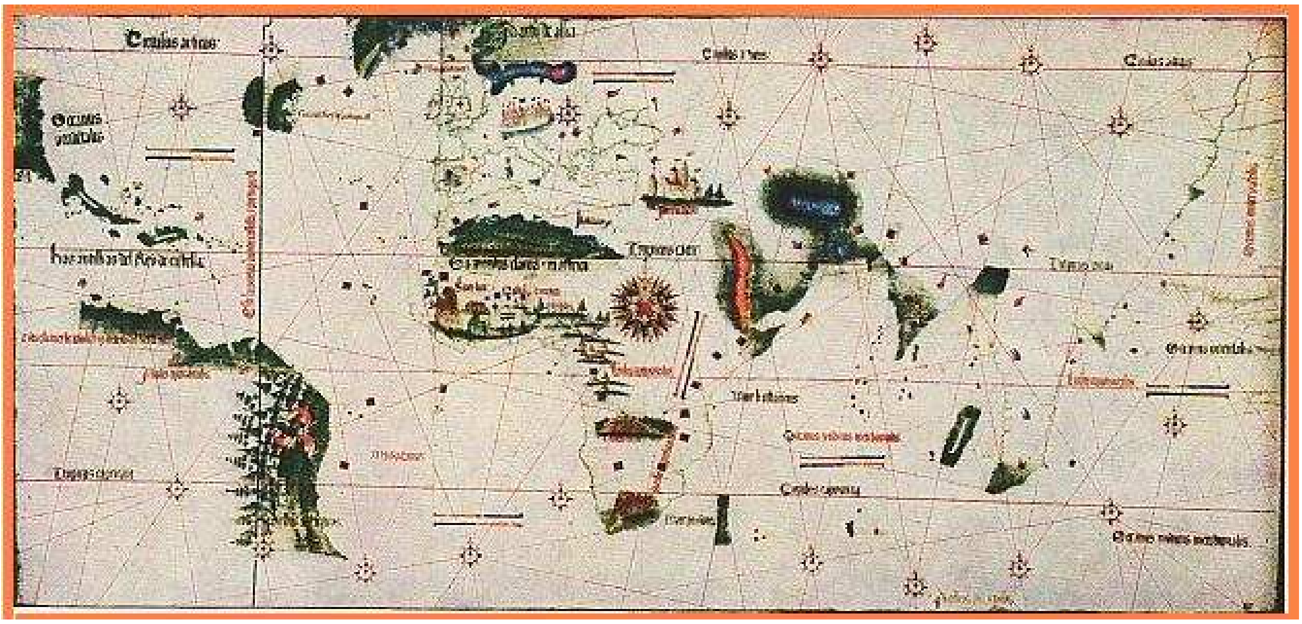
Al Idrisi's map



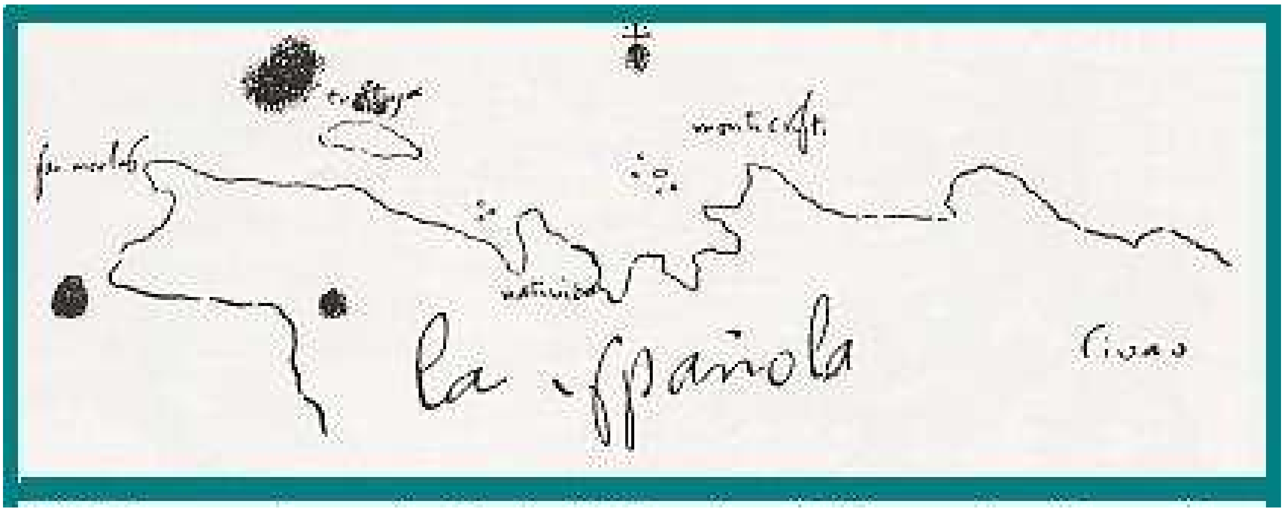
Ebstorf's map



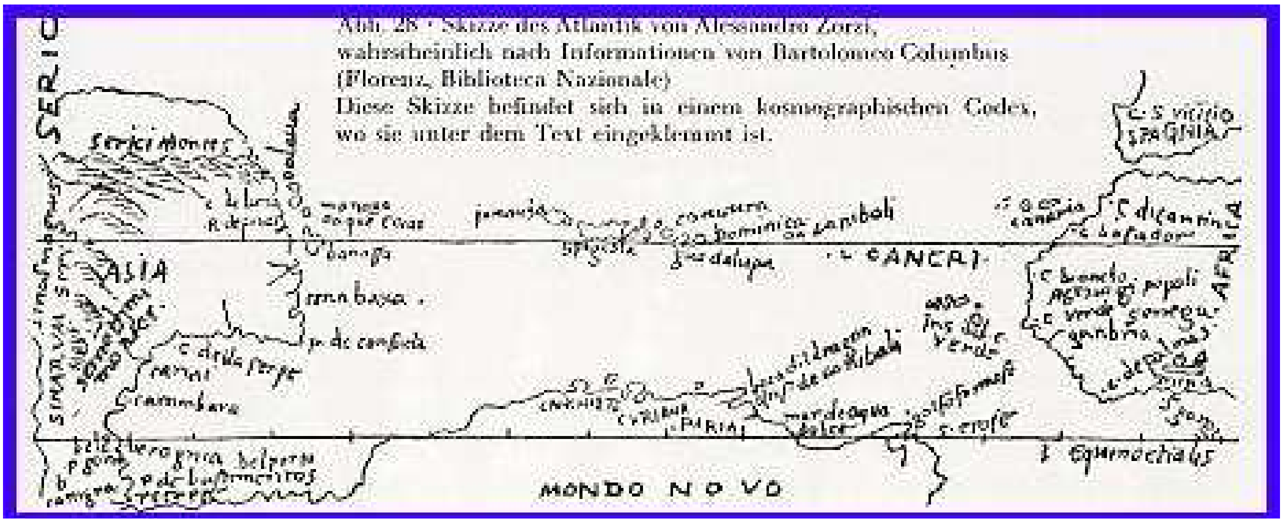
Medieval planisphere



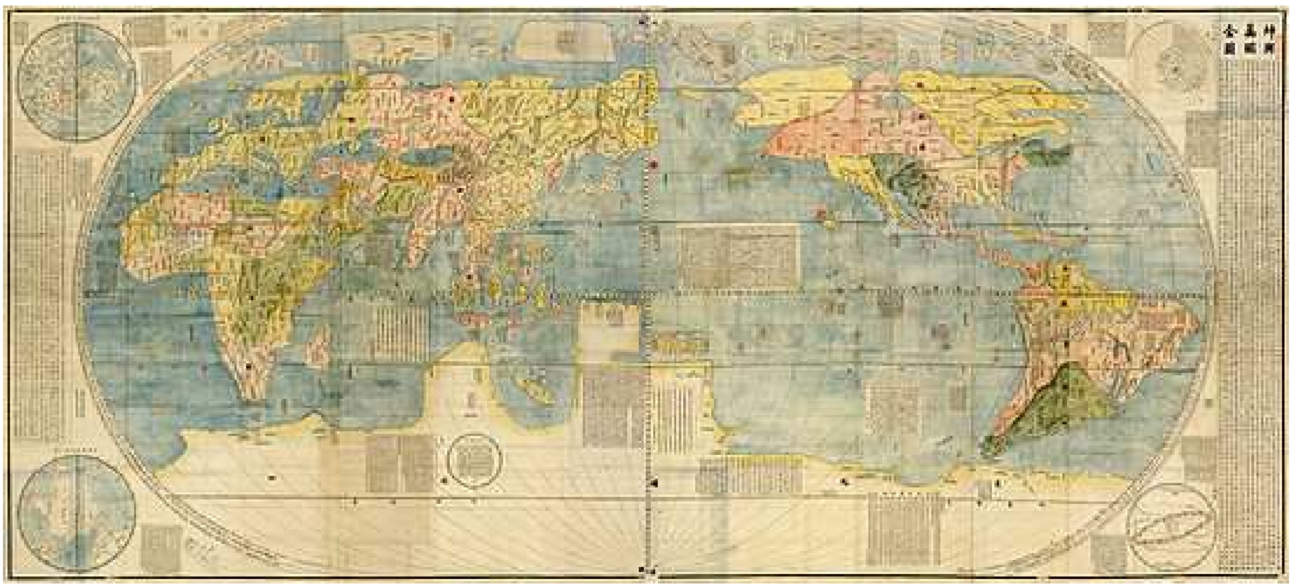
Renaissance planisphere



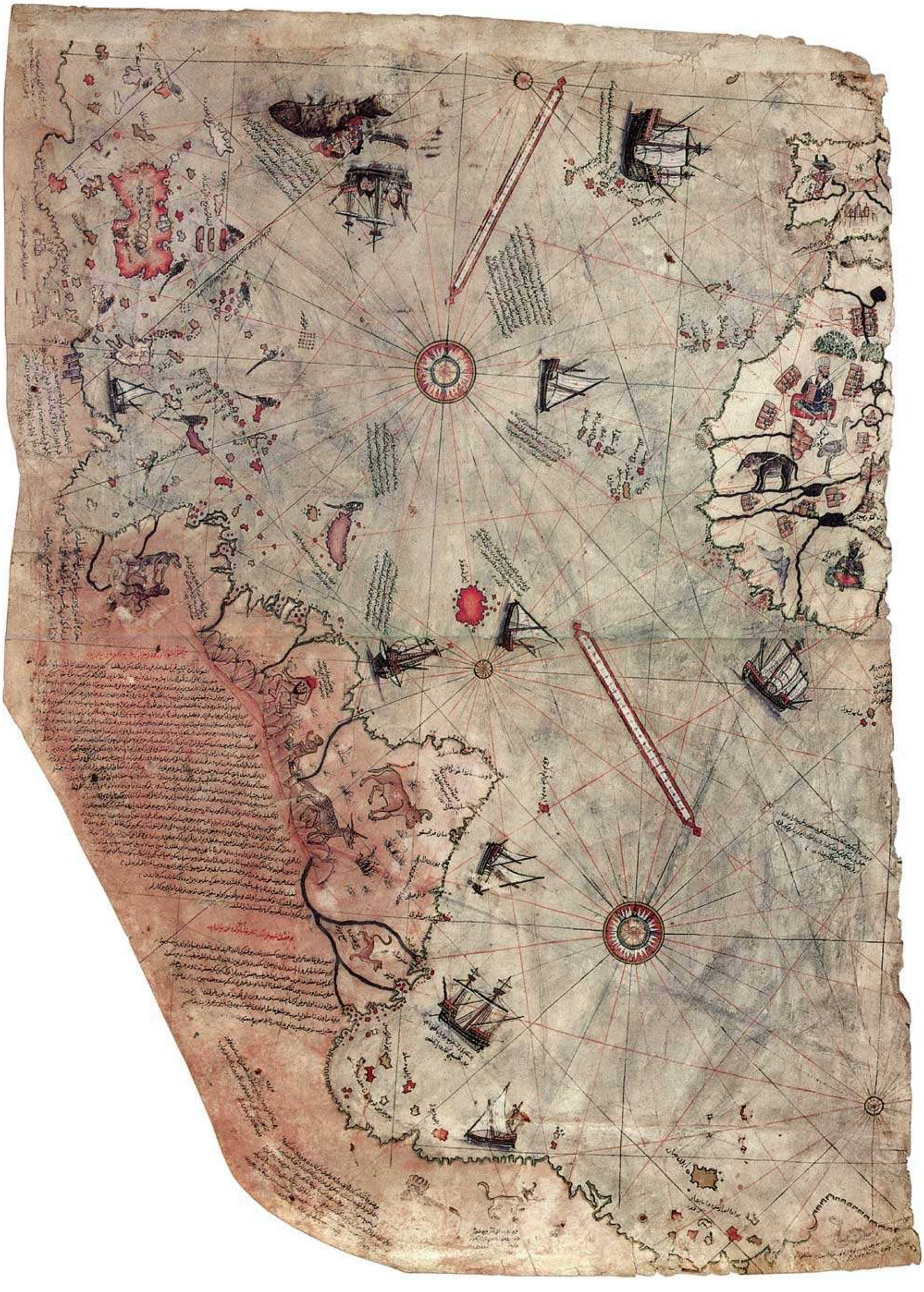
Hispaniola (Haiti) coast map



Novus “mundus” map



Matteo Ricci’s map



Piri Re'is' map

Mercator's map

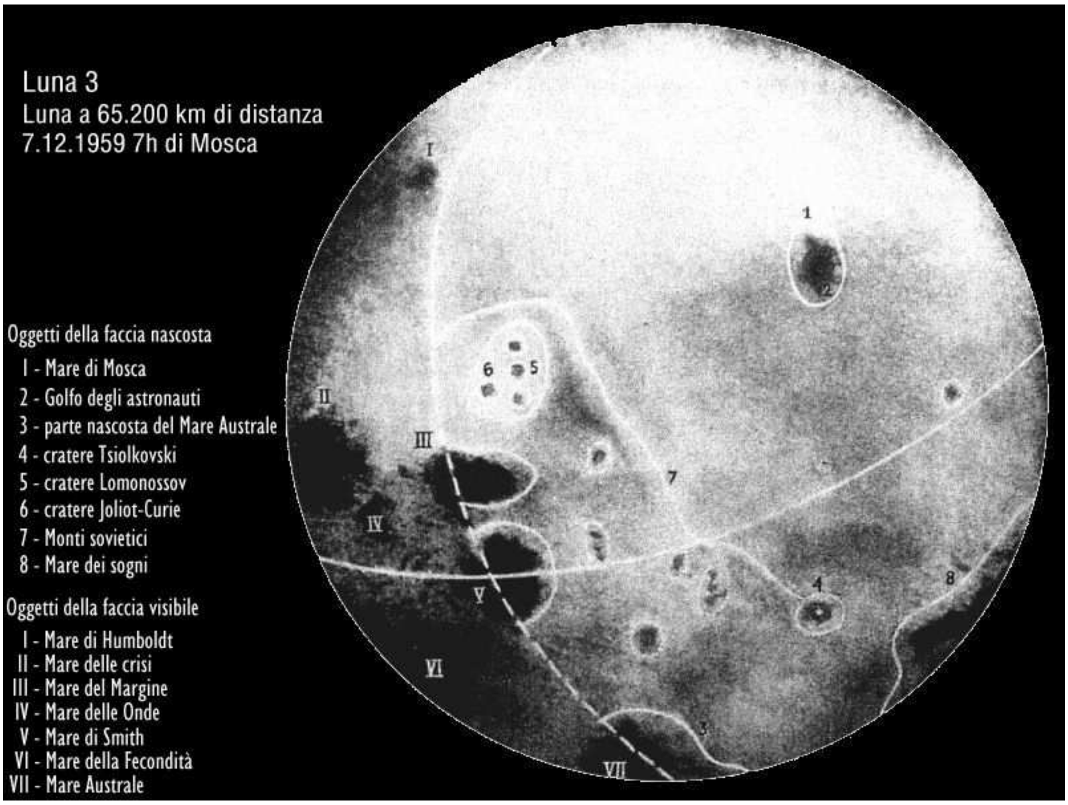
Mercator's planisphere

Tommaso Campanella's Sun city

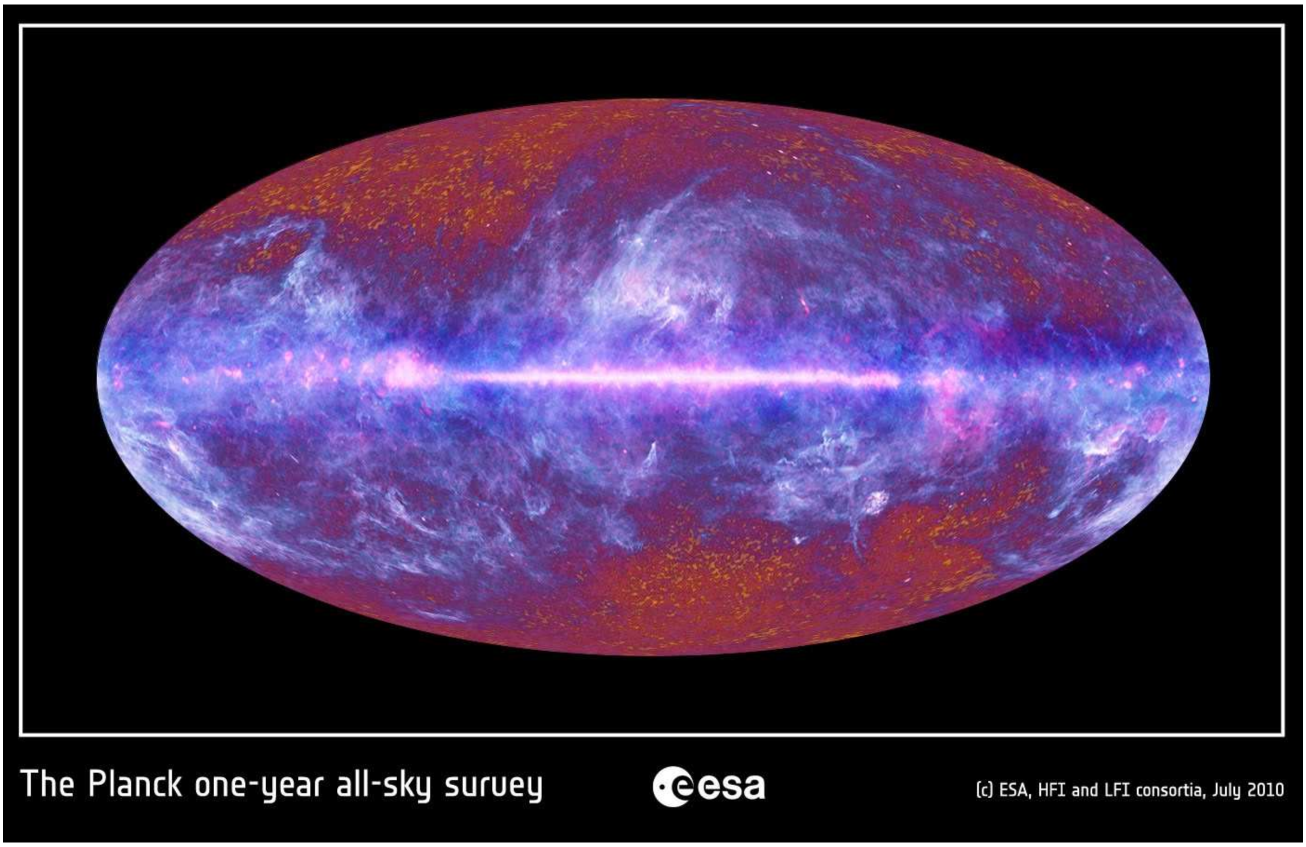
Nova Atlántida

Boscovich's Papal States map

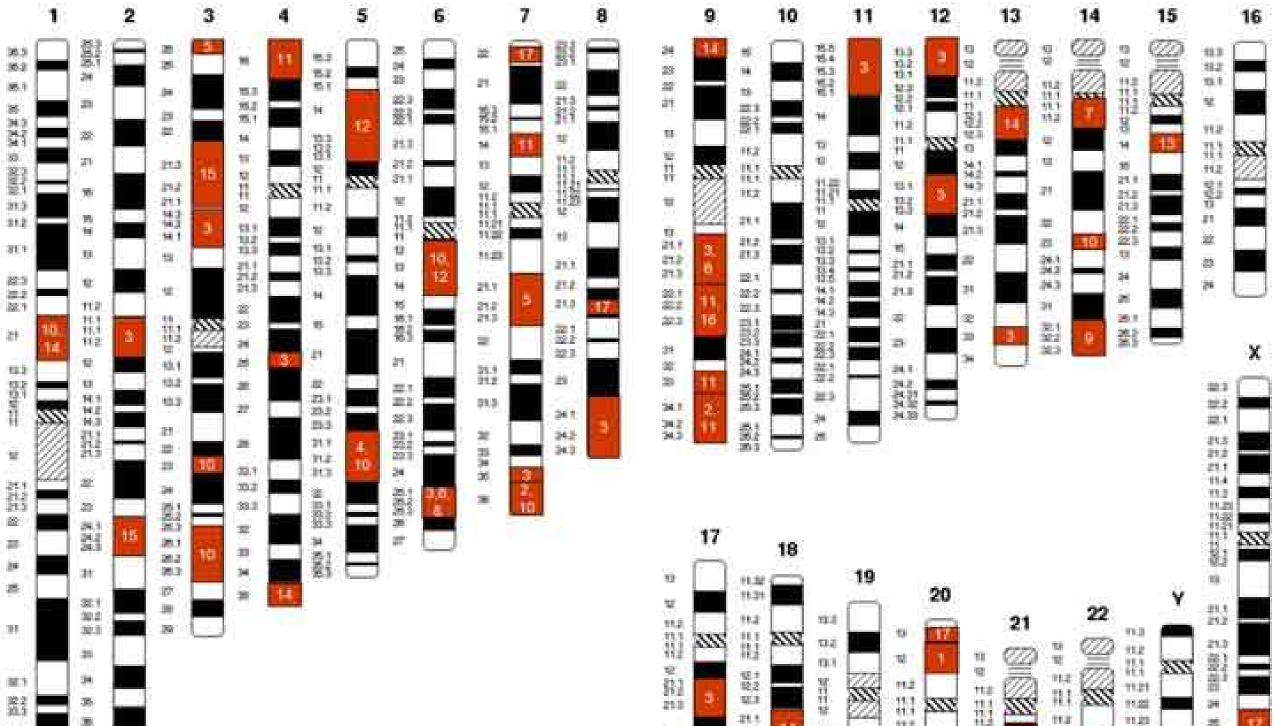
The Earth from space



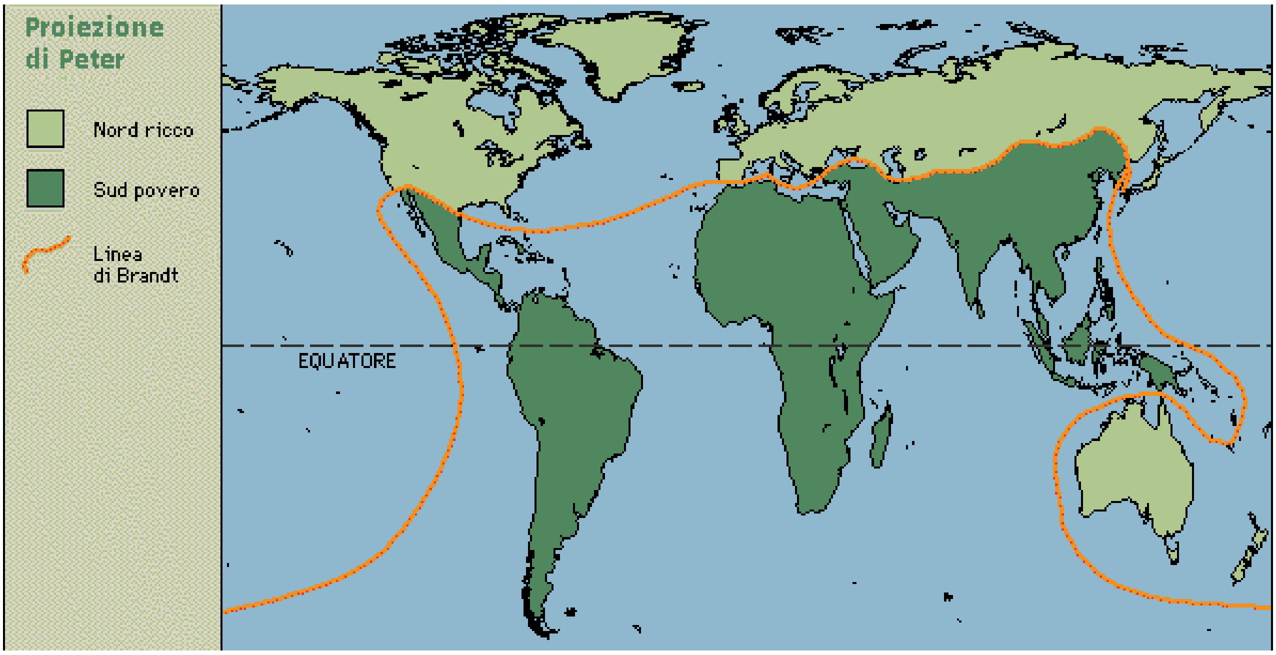
The Dark Side of the Moon



The actual known Universe map

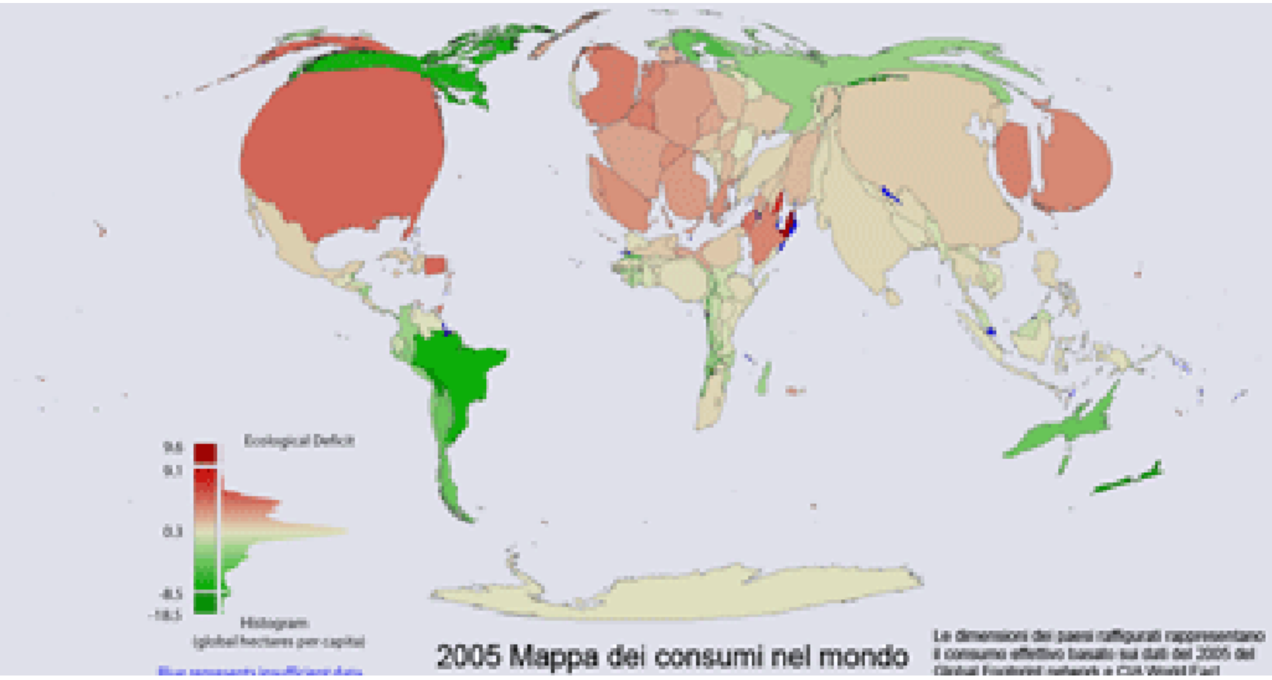


The human genome map

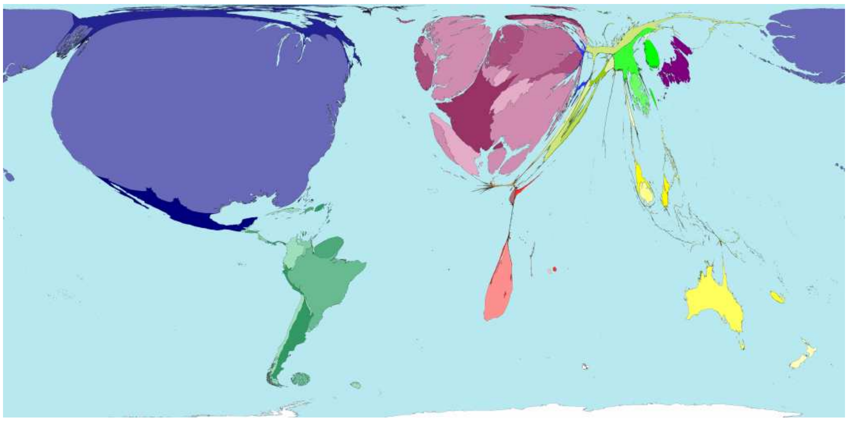


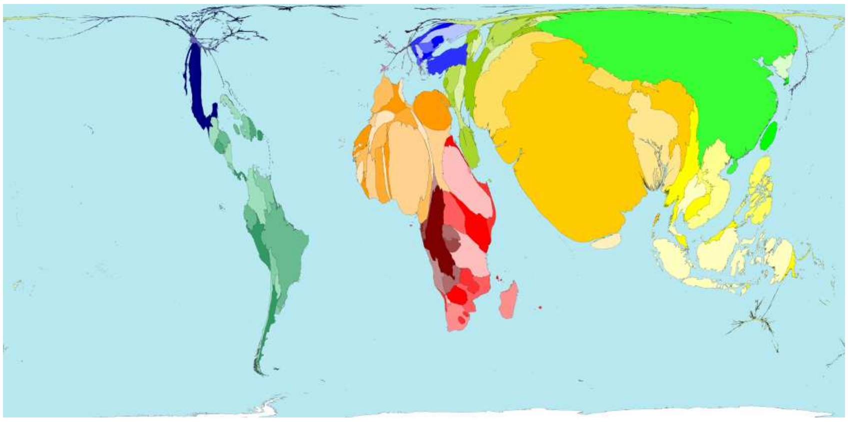
The Arno Peters' equivalent map

Thematic map of the world population

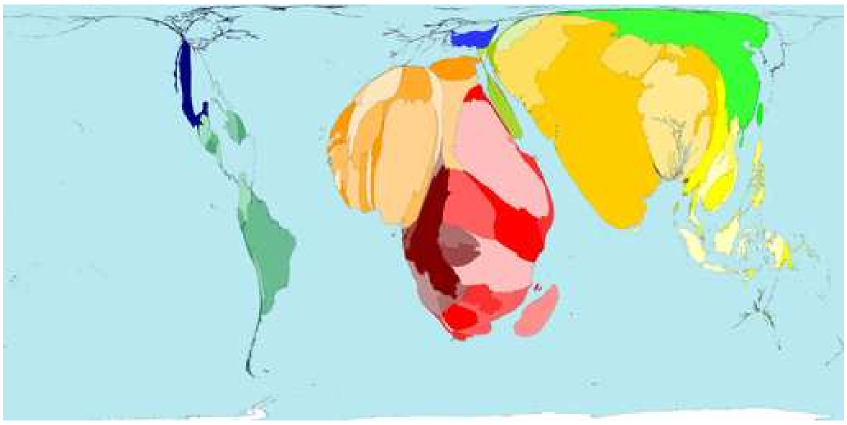


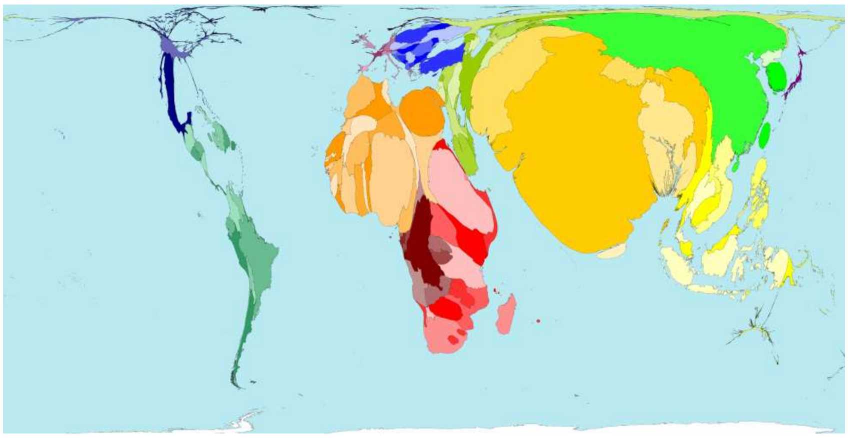
Thematic map of the global consumption





Thematic maps of families with more than 200 dollars/day and less than 10 dollars/day





Thematic map of the child labor

Thematic map of illiteracy

ESSENTIAL REFERENCES

1. Acheson J.D. (1952): Quality Control and Industrial Statistics. R.D. Irwin, Homewood (Illinois).
2. Agresti A. (1984): Analysis of Ordinal Categorical Data. Wiley & Sons, New York.
3. Andersen E.B. (1980): Discrete Statistical Models with Social Science Applications. North Holland Publishing Company, Amsterdam.
4. Ballard D., Brown C.M. (1982): Computer Vision. Prentice Hall, Englewood Cliffs (New Jersy).
5. Barnett V. (1973): Comparative Statistical Inference. John Wiley & Sons, London.
6. Barnett V., Lewis T. (1978): Outliers in statistical data. J. Wiley & Sons, New York.
7. Barry A.B. (1978): Errors in Practical Measurement in Science, Engineering, and Technology. John Wiley & Sons, New York.
8. Bartelme N. (1995): Geoinformatik – Modelle Strukturen Funktionen. Springer, New York.
9. Bartles R., J. Beatty and B.Barsky (1987): An introduction to the Serpent user Interface Management System. Software Engineering Institute. Carnegie–Mellon University, Pittsburgh (Pennsylvania).
10. Beyer W. H, (1966): CRC Handbook of Tables for Probability and Statistics. CRC Press, Boca Raton (Florida).
11. Bézier P. (1974): Mathematical and Practical Possibilities of UNISURF. Computer Aided Geometric Design, Academic Press, New York.
12. Bjerhammar A. (1973): Theory of Errors and Generalized Matrix Inverses. Elsevier Scientific Publishing Company, Amsterdam. [New York.
13. Blachut T.J., Chrzanowsky A., Saastamoinen J.H. (1979): Urban Surveying and Mapping. Springer,
14. Bloomfield P., Steiger W.L. (1983): Least Absolute Deviations. Birkhäuser, Boston.
15. Bonford B.G. (1952): Geodesy. Oxford University Press, London. [Boston.
16. Buchanan B. (1983): Constructing an expert system. Building Expert Systems. Addison Wesley,
17. Bury K.V. (1975): Statistical Models in Applied Science. Wiley & Sons, New York.
18. Chiswick B.R., Chiswick S.J. (1975): Statistics and Econometrics. A Problem–solving Text. University Park Press, Baltimore.
19. Cochran W.G. (1977): Sampling Techniques. John Wiley & Sons, New York.
20. Cox D.R., Hinkley D.V. (1974): Theoretical Statistics. Chapman and Hall, London. [Jersy).
21. Cramer H. (1946): Mathematical Methods of Statistics. Princeton University Press, Princeton (New
22. Cressie N.A. (1991): Statistics for Spatial Data. John Wiley and Sons, New York.
23. David F.N. (1954): Tables of The Ordinates and Probability Integral of the Distribution of the Correlation Coefficient in Small Samples. Cambridge University Press, Cambridge.
24. Davies O.L. (1960): The Design and Analysis of Industrial Experiments. Oliver and Boyd, London.
25. Davies O.L. (1961): Statistical Methods in Research and Production. Oliver and Boyd, London.
26. De Berg M., Cheong O., Van Kreveld M., Overmars M. (2008): Computational Geometry - Algorithms and Applications. Springer, New York.
27. Doebelin E.O. (1975): Measurement Systems. Application and Design. McGraw-Hill, New York.
28. Draper N.R., Smith H. (1981): Applied Regression Analysis. Wiley and Sons, New York.
29. Dwyer P.S. (1951): Linear Computation. John Wiley & Sons, New York.
30. Edberg E. (1976): Matrix computation. In: *Lectures given in the Academic Training Program of CERN* 1975–1976, CERN 76–22, Genève. [York.
31. Federer W.T. (1973): Statistics and Society. Data Collection and I nterpretation. M. Dekker, New
32. Fischer F.E. (1973): Fundamental Statistical Concepts. Canfield Press, San Francisco.
33. Foerstner W. (1992): Robust Computer Vision. Wichmann, Karisruhe.
34. Gentle J.E., Kennedy W.J. (1980): Statistical computing. Dekker, Basel.
35. George A., Liu J.W.H. (1981): Computer solution of large sparse positive definite systems. Prentice Hall, Englewood Cliffs (NJ).
36. Golub G.H., van Loan C.F. (1983): Matrix computation. North Oxford Academic, Oxford.
37. Grafarend E.W. et al. (1979): Optimierung Geodätischer Messoperationen. Wichmann, Kalsruhe.
38. Grimshaw D.J. (1994): Bringing Geographical Information Systems in Business. Longman, London.
39. Grimson W.E.L. (1981): Theory of human stereo vision. Mc Graw-Hili, New York. [New York.
40. Guttman I., Wilks S.S., Hunter S. J. (1965): Introductory Engineering Statistics, John Wiley & Sons,
41. Hageman L.A., Young D.M. (1981): Applied Iterative Methods. Academic Press, New York.
42. Hampel F.R., Ronchetti E.M., Rousseeuw P.J., Stahel W.A. (1986): Robust Statistics. The Approach Based on Influence Functions. Wiley & Sons, New York.
43. Hawkins D.M. (1980): Identification of outliers. Chapman and Hall, London.
44. Hawkins D.M. (1982): Topics in Applied Multivariate Analysis. Cambridge University Press, London.
45. Heiskanen W.A., Moritz H. (1967): Physical Geodesy. Freeman and Company, San Francisco.
46. Huber P.J. (1981): Robust Statistics. Wiley & Sons, New York.
47. Jobson J.D. (1991/2): Applied Multivariate Data Analysis: vol. I & II. Springer, New York.
48. Kaufman L.,Rousseeuw P.J. (1990): Finding Groups in Data. Wiley & Sons, New York.

49. Kendall M. (1980): Multivariate Analysis. Charles Griffin & Company LTD, London.
50. Koch K.R. (1987): Parameter Estimation and Hypothesis Testing in Linear Models. Springer, Berlin.
51. Kramer B. (1988): The Art of Measurement. Metrology in Fundamental and Applied Physics. VCH, Weinheim (Rhein-Neckar).
52. Kraus K., Waldhäusl P. (1994): Photogrammetrie. Dümmler, Bonn.
53. Lambeck K. (1988): Geographical Geodesy – The Slow Deformation of the Earth. Oxford University Press, Oxford. [London.
54. Laurini R., Thompson D. (1996): Fundamentals of Spatial Information Systems, Academic Press,
55. Lauritzen S.L. (1998): Graphical Models. Clarendon Press, Oxford.
56. Lawson G.L., Handson R.J. (1974): Solving least squares problems. Prentice Hall, *Series in Automatic Computation*, Englewood Cliffs (New Jersy).
57. Maguire D.J. (1991): An Overview and Definition of GIS: Geographical Information Systems. Longman, London.
58. Marr D. (1982): Vision. Freeman, San Francisco.
59. McPherson G. (1990): Statistics in Scientific Investigation. Springer, New York.
60. Meissl P. (1982): Least Squares Adjustment a Modern Approach. Mitteilungen der geodätischen Institute der Technischen Universität Graz, Folge 43, Graz.
61. Middlebroks E.J. (1976): Statistical Calculation, How to Solve Statistical Problems. Ann Arbor Science, Ann Arbor (Michigan).
62. Mikhail E.M., Ackermann F. (1976): Observations and Least Squares. IEP–A Dun–Donnelley Publisher, New York.
63. Modi J.J. (1988): Paraliel algorithms and matrix computations. Oxford Univ. Press, London.
64. Mood A.M., Graybill F.A., Boes D.C. (1974): Introduction to the Theory of Statistics. Mc Graw–Hill Book Company, London.
65. Moritz H. (1980): Advanced Physical Geodesy. Wichmann, Karlsruhe.
66. Mortenson M. E. (1997): Geometric Modeling. John Wiley and Sons Inc., New York. [York.
67. Papoulis A. (1965): Probability Random Variables and Stochastic Processes. McGraw–Hill, New
68. Pavlidis S. (1977): Structurai pattern recognition. Springer, New York.
69. Pfeiffer P.E. (1978): Concepts of Probability Theory. Dover Publications, New York. [York.
70. Preparata F.P., Shamos M.I. (1985): Computational Geometry – an Introduction. Springer, New
71. Rao C.R. (1965): Linear Statistica] Inference and Its Applications. Wiley & Sons, New York.
72. Richardus P., Adler R.K. (1972): Map Projections. North-Holland Publishing Company, Amsterdam.
73. Richardus P., Allam J.S. (1977): Project Surveying. North-Holland Publishing Company, Amsterdam.
74. Rousseeuw P.J, Leroy A.M. (1987): Robust Regression and Outlier Detection. Wiley & Sons, New York.
75. Sachs L. (1984): Applied Statistics. A handbook of Techniques. Springer, New York.
76. Savage L.J. (1954): The Foundations of Statistics. John Wiley & Sons, New York.
77. Simon H. A. (1960): The New Science of Management Decision. Harper & Row, New York.
78. Strang G. (1988): Linear algebra and its applications. Harcourt Brace Janovich, San Diego.
79. Tanner M.A. (1996): Tools for Statistical Inference. Springer, New York.
80. Telford W.M., Geldart L.P., Sheriff R.E., Keys D.A. (1976): Applied Geophysics. Cambridge University Press, Cambridge.
81. Tewarson R. P. (1972): Sparse matrices. Academic Press, *Mathematics in Science and Engineering*, n. 99, New York. [London.
82. Thompson E.H. (1969): An Introduction to be Algebra of Matrices With Some Applications. A. Hilger,
83. Thurston W. (1997): Three-dimensional Geometry and Topology. Princeton University Press, Princeton (New Jersy).
84. Torge W. (1980): Geodesy. De Gruyter, Berlin. [Amsterdam.
85. Vaníček P., Krakiwsky E. (1982): Geodesy: the Concepts. North-Holland Publishing Company,
86. Varga M.S. (1962): Matrix iterative analysis. Prentice Hall, Englewood Cliffs (Nw Jersy).
87. Wald A. (1947): Sequential Analysis. Wiley & Sons, New York.
88. Watson G.N. (1948): A Treatise on the Theory of Bessel Functions. Cambridge University Press, Cambridge, Mac Millan Company, New York.
89. Weeks J. (1985): The Shape of Space, Dekker, New York.
90. Whittle P. (1990): Risk–sensitive Optimal Control. John Wiley & Sons, Chichester (West Sussex).

TABLE OF CONTENTS

OVERVIEW **1**

PART I – ELEMENTARY STATISTICS **2**

PROBABILITY THEORY **2**

Definition of probability **2**

"Geometric" or Laplace probability **2**
"Frequency" or Von Mises probability **2**
"Axioms" or Kolmogorov probability **2**

Theorems of probability **2**

Sum **2**
Composed **2**
Bayes **2**

DESCRIPTIVE STATISTICS **2**

Simple statistics variables **2**

Position (or center) indices **3**

- ❑ Mean **3**
- ❑ Median **3**
- ❑ Mode **3**

Dispersion's indices **3**

- ❑ Variance **3**
- ❑ M.A.V. **3**
- ❑ m.a.v. **3**
- ❑ Amplitude **3**

Shape Indices **3**

- ❑ Asymmetry **3**
- ❑ Kurtosis **3**

Chebychev's theorem **3**

Means and mutual variability **4**

2D statistical variables **4**

Connection **5**

- ❑ Bonferroni unilateral indices **5**
- ❑ Bonferroni bilateral indices **5**

Correlation (linear dependence) **5**

- ❑ Linear correlation coefficient **6**
- ❑ Regression lines **6**
- ❑ Robust regression lines **6**

Regression 6

- ❑ Orthogonal variance decomposition theorem 7
- ❑ Pearson unilateral indices (max dependence index **7**
- ❑ Pearson bilateral index (max dependence index) **7**
- ❑ Robust regression **7**

TRANSFORMATION OF RANDOM VARIABLES **7**

Variance and covariance propagation law **9**

EXAMPLES OF RANDOM VARIABLES **9**

Normal random variable (for measurement random errors) **9**
Uniform random variable (for small segments) **10**
Log - normal random variable (for rare events) **10**

Central Limit Theorem and Law of Large Numbers **10**

STATISTICAL INFERENCE **10**

Parametric test **10**

Large sample mean test and confidence bond **11**
Power curve of normal distribution **11**
Normal sample tests and confidence bounds **12**

Distribution – free (non – parametric) tests **12**

Goodness of fit Test **12**
Independence Test **13**
Test of Kolmogorov-Smirnov **13**
Independence test by the test of Kolmogorov-Smirnov **13**
Test of Pearson et al. for normality **13**
Test of Mann-Whitney **14**
Test of Siegel-Tuckey **14**
Test of sign(for central values) **14**
Sign test (for the dispersion’s values) **15**
Test of Spearman **15**

Welch and Tukey test of mean comparison of normal samples with different variances **16**

PART II – ESTIMATION THEORY **17**

Condition equations **17**
Observation equations (with parameters) **18**
Observation equations (with parameters and constraints) **19**
Observation equations (with parameters and over-weighted pseudo-observations) **19**
Observation equations (with parameters and generic pseudo-observations) **20**

Complements **21**

Direct constraints and pseudo-observations **21**
Slack parameters **21**
Supplementary observations **21**

Non-linear problems **22**

Functional model (Newton-Fourier method) **22**

- ❑ Tikhonov regularization **22**
- ❑ Levenberg-Marquardt algorithm **22**

Stochastic model (Gauss-Helmert model for weight reproduction) **22**
Norm **23**

- ❑ Huber estimator (M estimators: minimax) **23**
- ❑ Hampel estimator (S estimators) **24**
- ❑ Auxiliary weight function **24**

Numerical tests **24**

Conditioning (of the parameters) **24**

- ❑ Singular value decomposition **24**
- ❑ Condition number **25**
- ❑ Conditioning via reliability **25**

Reliability (of the observations) **25**

- ❑ Internal reliability **26**
- ❑ External reliability **26**

Appendix A – Robust estimators **26**

Estimator variance **26**
Influence function **26**
Rejection point **26**
Gross–error sensitivity **26**
Breakdown point **26**

- ❑ Least median of squares (LMedS) **26**
- ❑ Random sampling (RANSAC) **27**

Appendix B – Partitioned models **27**

Kalman filter **27**
Gauss blocking **28**

Appendix C – Sampling **29**

Logistic strategies **29**
Stratified sampling **31**

Appendix D – Design of experiments **31**

PART III – MULTIVARIATE ANALYSIS **32**

CLUSTER ANALYSIS **32**

MULTIPLE REGRESSION **33**

Orthogonal variance decomposition and error minimization **34**
Fisher tests **34**
Regression validation **35**

VARIANCE ANALYSIS (ANOVA) **35**

One way **35**
Two ways **35**
Two ways with interaction **35**
Fisher tests **35**

Three ways **36**
Three ways with one interaction **36**
Three ways with three interactions **36**
Three ways with three interactions and a multiple interaction **36**
Fisher tests **37**

SEQUENTIAL TESTS **37**

PART IV – INTERPOLATION AND APPROXIMATION **38**

FINITE METHOD INTERPOLATION **38**

One-dimension cubic case 38
Two-dimension cubic case 38

SPLINE INTERPOLATION 38

One-dimension linear case 38
One-dimension cubic case 38
Two-dimension linear case 39
Two-dimension cubic case 39
Interpolation and extrapolation 40

COVARIANCE ESTIMATION 41

Empirical estimation 41
Empirical optimization of the spacing 41
Theoretical models 42
Finite covariance functions 43
Space-temporal problems 43

COLLOCATION (FILTERING AND PREDICTION) 43

GENERALIZED LEAST SQUARES 44

KRIGING 45

Semi-variogram 45

Simple Kriging 45
Ordinary Kriging 45
Ordinary Co-Kriging 45
Universal Kriging 46

PART V – NUMERICAL METHODS 47

DIRECT METHODS 47

Gauss' elimination 47

Cholesky factorization 47
Forward and backward substitution 47
Inversion of the normal matrix 48

Sparse matrices 48
Variance propagation 49

ITERATIVE METHODS 50

Conjugate Gradient method 50
Preconditioning and ICCG methods 50

- Incomplete Cholesky factorization 50
- ICCG method 50

Approximate inversion 51

REGULAR STRUCTERES 51

Toeplitz matrix 51
Kronecker decomposition 51
Trench algorithm 52

SEQUENTIAL ALGORITHMS 52

Householder transformation **52**
Sequential Householder transformation **53**

Givens algorithm **53**
Sequential Cholesky factorization **54**
Matrix inversion lemma **54**
Block Gauss inversion **55**

GRAPH THEORY **55**

Reordering **55**
Dissection and reordering **56**

PART VI – EXERCISES **58**

TWO EXAMPLES OF DESCRIPTIVE STATISTICS **58**

One-dimension statistical variable **58**
Two-dimension statistical variable **59**

EXAMPLES OF NORMAL AND DISTRIBUTION-FREE TESTS **61**

Parametric tests according to independence and normality hypotheses **61**
Distribution-free tests **63**

TWO EXAMPLES OF LEAST SQUARES PROBLEMS **64**

Linear interpolation in a space domain **63**
Cubic polynomial interpolation in a time domain **65**

THREE EXAMPLES OF GEODETIC LEAST SQUARES PROBLEMS **67**

Adjustment of two lattice structures **67**

- Finite differences of first order 67
- Finite differences of second order 68

Surface reconstruction **69**

PART VII – DATA PROCESSING MEETS HUMAN SCIENCES **71**

THE ANCIENT WORLD **71**

Signs and symbols **71**
Origin of writing **71**

Ancient Mathematics **72**
Astronomy **72**
Travel and cartography **72**
Ancient Philosophy **72**
Ancient Medicine **73**

THE MIDDLE AGES AND MODERN AGE **73**

From the fall of the ancient world to modern age **73**
Travels **74**

Philosophy **74**
Mathematics and Mathematical physic **75**
Music **75**

Sciences **75**

- Astronomy **75**
- Geodesy e cartography **75**
- Physics **76**
- Chemistry **76**
- Biology **76**
- Medicine **76**
- Linguistic **76**

CONTEMPORARY AGE AND PRESENT WORLD **76**

Philosophy **76**
Logics **77**
Mathematics **77**
Physics **77**
Informatics **77**
Linguistic **78**

Science **78**

- Biology **78**
- Medicine **78**
- Psychology **78**

The conquest of space (locomotion and energies) **78**

Literature Art and Music **79**
Pedagogy and Ethics **79**
Time and Place of Utopia **79**

A map collection along the history path **80**

- Herodotus' Ecumene map **80**
- Ptolomy's Ecumene map **80**
- Jan Van der Straet, Dante Alighieri's Hell **81**
- Ambrosius Holbein, Thomas More's Utopia **81**
- Isidore of Siville's map **82**
- Ebstorf's map **82**
- Medieval planisphere **83**
- Renaissance planisphere **83**
- Hispaniola (Haiti) coast map **84**
- Novus "mundus" map **84**
- Matteo Ricci's map **84**
- Piri Re'is' map **85**
- Mercator's map **86**
- Mercator's planisphere **86**
- Tommaso Campanella's Sun city **87**
- Nova Atlántida **87**
- Boscovich's Papal States map **88**
- The Earth from space **89**
- The Dark Side of the Moon **89**
- The actual known Universe map **90**
- The human genome map **90**
- The Arno Peters' equivalent map **91**
- Thematic map of the world population **91**
- Thematic map of the global consumption **91**
- Thematic maps of families with more than 200 dollars/day and less than 10 dollars/day **92**
- Thematic map of the child labor **92**
- Thematic map of illiteracy **92**

ESSENTIAL REFERENCES **93**

TABLE OF CONTENTS **95**