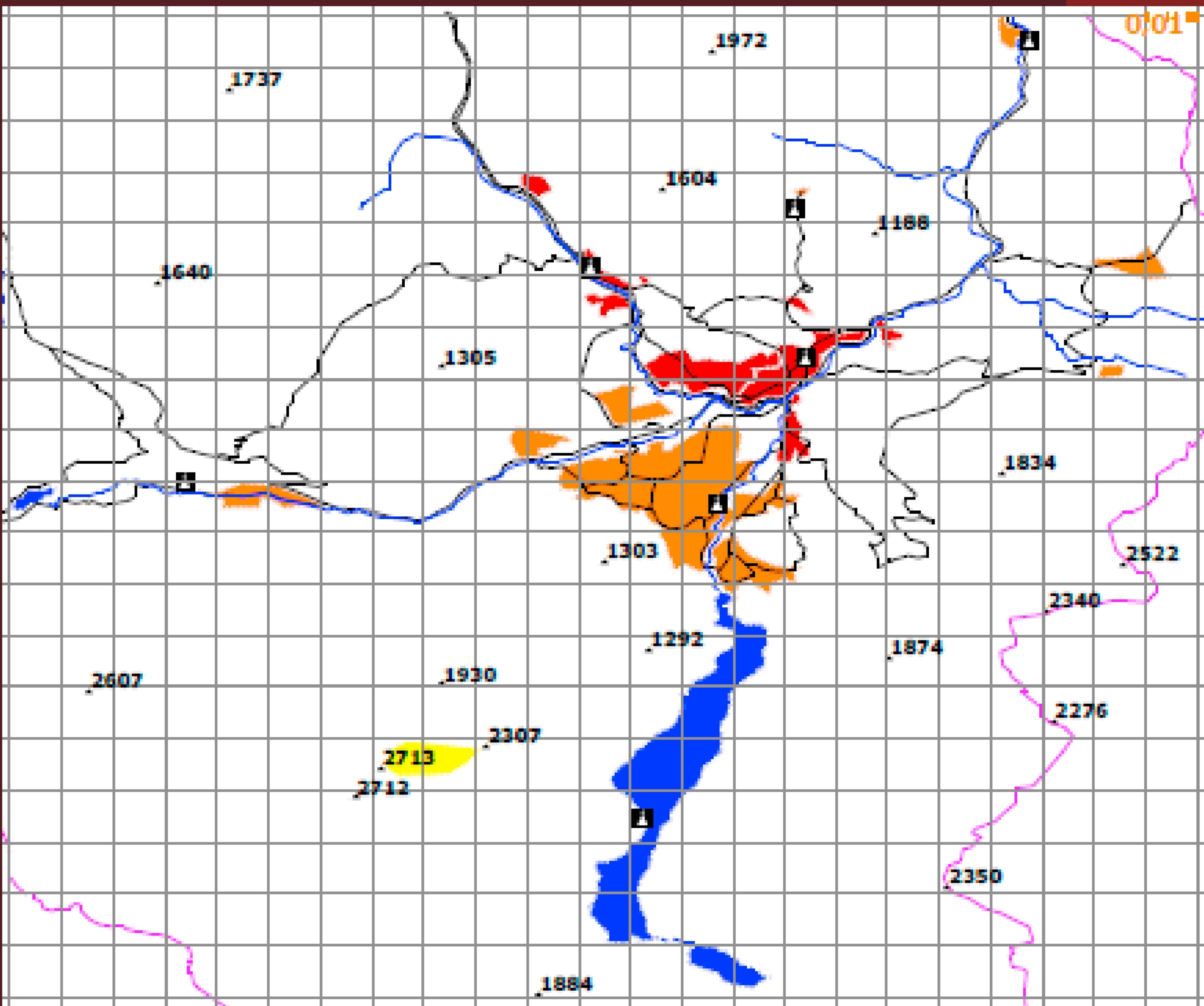
Markus Penzkofer

Geoinformatics



from basic to specialized knowledge

Geoinformatics
basic and specialized knowledge

Statistics and Adjustment Computation

Markus Penzkofer

First Edition
December 2016

Geoinformatics - from basic to specialized knowledge including instruction to the Desktop GIS MensorGIS

Markus Penzkofer

November 2016

Image on the cover:
Map extract around Königssee, Bavaria, Germany,
generated using the printing function of MensorGIS

Herstellung und Verlag:
BoD - Books on Demand, Norderstedt
ISBN 978-3-7431-0803-5

Contents

List of Figures

Chapter 1

Introduction

1.1 Applications of Geoinformatics

Geoinformatics has many different fields of application, which are nowadays even known by the general public. Pure expert systems are now the exception, but behind every user interface of course there is still a core, which has to be administered by a professional administrator.
Usually the fields of application of GIS Applications are split up as follows:

- Branch Information Systems, e.g. Providers of Base Data, Environmental Information Systems, Energy- and Water Service Providers, Fleet Management
- Public Information Systems, e.g. Routing, Google Earth, Location Based Services, Geodata Portals

1.2 Aim of this Paper

In this manuscript some theoretical background and interrelations of Geographic Information Systems shall be assembled. Also the most important formulae of Mathematics, Surveying, Geodesy and Cartography concerning Geoinformatics will be presented in short. Besides these provinces of course Informatics builds a very important part of GIS Applications. In a final chapter all basic knowledge will be combined to a short overview over Geoinformatics, where especially modelling, extensions, specializations and analysis of Geoinformatics are shown.
In the last chapter functionality and source code of the Open Source GIS **MensorGIS** are discussed regarding the basics explained before. The aim of this GIS Project is not its operative use, but the exemplary implementation of algorithms. MensorGIS is available for free under: `https://sourceforge.net/projects/mensorgis/` .
This paper can only touch each single province with basic topics; if some things are missed, technical literature should be taken to look them up (see Bibliography in Appendix A). The Book should be read by students, also by lateral entrants, approaching advanced lectures with the focus on Geoinformatics.

The next chapter gives an overview over the individual disciplines, which are needed when conceiving, programming and using a GIS. A relevant aspect in these chapters is, what to keep in mind when implementing the matters.
Important terms are printed in bold face, additional terms and references between the chapters in italic characters. For better readability important terms are written in uppercase letters.

Chapter 2

Overview over the Basic Knowledge of Geoinformatics

When talking about GIS Applications, often knowledge of Mathematics, Geodesy and Informatics is mentioned. But when thinking of the origin of the GIS Data, as well some of the measurement techniques of Surveying before the age of GPS as well as the Photogrammetric computation of Orthophotos and Digital Terrain Models have to be mentioned. In addition some important methods of Cartography are also needed in GIS Applications. The province of Surveying and Geodesy brings in the traditional knowledge, Informatics makes available the basis for the utilization of new technologies in information processing of spatial data.

2.1 Mathematics

Mathematics builds the basis for many Engineering Sciences, but also for the theoretical and practical applications of Informatics. Regarding Geoinformatics, it influences beginning with Surveying and Geodesy, next Cartography and Photogrammetry finally also Informatics and Geoinformatics.
The basic geodetic applications mainly regard **Trigonometry**, **Geometry** and **Matrix Computation**. Cartography and Photogrammetry need the same basic knowledge, additionally also **Interpolation Methods**. In Informatics concerning programming, the **Representation of Numbers** and the **Boolean Logic** play a role. A specific GIS Analysis is based on the **Graph Theory** as part of mathematical Topology.

2.2 Surveying

Surveying covers all calculations, which can be executed by assumption of a plane Reference Surface. To do so a local plane Coordinate System is used or calculation is done using Coordinates retrieved by a Projection. Elevation values are typically measured separately from Planimetric Survey (exception is GPS).
The most practical **Measurement Data** come from Surveying, when concerning that measurement by GPS nowadays is "normal" Surveying practice.

2.3 Geodesy

Geodesy is based on the definition of a Reference Surface, which is curved in opposition to Surveying, i.e. in former times a Sphere, in the young history an Ellipsoid or Geoid to represent the Earth. Geodesy defines country- or world-wide **Reference Systems**. Besides this, Geodesy deals with the measurement and definition of the Gravity Field of the Earth (concerning GIS, here no further thinking).
A method, which is used as well in Surveying, Geodesy as well as in Photogrammetry, is computation by Adjustment Methods (see also the second part of this book).

2.4 Cartography and Photogrammetry

By the methods of **Cartography** a systematical representation (Mapping) of measured and calculated objects can be done. Two parts of Cartography are fundamental for Geoinformatics: On the one hand the mathematical calculation methods for projecting these objects onto the plane (**Map Projection**), on the other hand the graphical processing and arrangement of the objects in Geometry and Appearance (e.g. **Generalization**, Color, Hatching, **Thematic Representation**).
Photogrammetry derives three-dimensional Coordinates from two-dimensional Image Data. In this way also data can be produced, which can be used for Mapping and/or as input to a GIS. Like Surveying, Photogrammetry therefore is a provider of Data for GIS Applications, and Cartography can use these and also the data measured by terrestrial Surveying.

2.5 Informatics

Informatics nowadays provides Computational and Data Storage Tools. From line-wise input into Computer Algebra Systems to creation of standalone programs with Graphic User Interface (GUI) there are many possibilities of Informatics, to automatize calculations.
Standalone Programs usually offer additionally Import, Export and permanent Storage of Data in a **Database**. This requires a thought-out **Data Modelling**. An essential component of contemporary GIS Applications is the possibility of graphical Visualization of data. Here **Computer Graphics** plays an important role. In the age of the Internet **Distributed Systems** have a large circulation.

2.6 Geoinformatics

Geoinformatics combines many of the basics mentioned so long. The kind of **Data Storage** and **Data Modelling** as well as the possibilities of **Visualization** and **Analysis** make GIS Applications special. With regard to Informatics and also the other classical provinces, there are many extensions and specializations. These topics will be mentioned in the chapter on *Geoinformatics*.

Chapter 3

Basics of Mathematics

Mathematics of course is much to extensive, to be presented here in a complete way. In the following parts of the Mathematics chapter only the fundamental knowledge concerning Surveying, Cartography and Geoinformatics shall be captured.

3.1 Trigonometry

The **Trigonometric Functions** Sine, Cosine, Tangent, their Inverse and other trigonometric Functions appear as well in Surveying/Geodesy (e.g. using Transformations) as well as in Cartography (e.g. using Projections) very often.
When programming, to keep in mind is, that the most Programming Languages demand their functional parameters in Radians. **Radians** is the representation in multiplies of π and is written in formulae as rad.

3.1.1 The Radians

The Radians is the Length of the Arc, which lies between the sides of the central angle of the Unit Circle (Radius = 1).
The Factor of Conversion between Degrees (Mathematics) and Radians is:

$$\begin{aligned}
\rho &= \frac{2 \cdot 1 \cdot \pi}{360^o} \\
\rho &= \frac{\pi}{180^o} \\
\rho &= 0.01745329251994\,[rad/^o]
\end{aligned}$$

For Grad (Gon) the following Factor of Conversion must be used:

$$\begin{aligned}
\rho_g &= \frac{2 \cdot 1 \cdot \pi}{400^g} \\
\rho_g &= \frac{\pi}{200^g} \\
\rho_g &= 0.01570796326795\,[rad/^g]
\end{aligned}$$

Mentioning this seems banal, but the wrong or forgotten Factor of Conversion often leads to annoying results in long calculations (programs).

3.1.2 The Right-angled Triangle

Formulae to solve the Right-angled Triangle appear very often in Surveying. The side opposite to the Right Angle is called **Hypotenuse**, the both other sides are called **Catheti** or **Legs**. Therefore briefly the Definitions of the fundamental Trigonometric Functions in the Right-angled Triangle as a rule of thumb:

- Sine = Opposite Leg to Hypotenuse
- Cosine = Adjacent Leg to Hypotenuse
- Tangent = Opposite Leg to Adjacent Leg

Further mathematical relations, like e.g. trigonometric Identities, may be taken from a formulary!

3.2 Statistics

Also Statistics is a big part of Mathematics; here only the fundamental methods and segments are outlined (more details in the second part of this book!).

3.2.1 Descriptive Statistics

Descriptive Statistics calculates empirical Values from a Set of Data, which characterize these data. The result are e.g. **Frequency Distributions** (*Histograms*), **Measures of Central Tendency** (*arithmetic Mean*, *geometric Mean*, *weighted Mean*, *Median*) and **Measures of Dispersion** (*Variance*, *Standard Deviation*).
Arithmetic Mean $\hat{x}$ and Variance σ^2 resp. Standard Deviation σ are the Values used most often, here are their formulae (for n values x_i):

$$\hat{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \qquad \text{Arithmetic Mean}$$

$$\sigma^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (\hat{x} - x_i)^2 \qquad \text{Variance}$$

$$\sigma = \sqrt{\sigma^2} \qquad \text{Standard Deviation}$$

The **Median** q_{50} is a Mean Value, which is less strongly influenced by Outliers than the other Mean concepts. The Numbers x_i to process have to be sorted at first and the "mean" Index has to be built, then there are two cases:

$$i_q = \mathrm{FIX}\left(\frac{\mathrm{n}}{2}\right) \qquad \text{Mean-Index}$$

$$q_{50}^b = x_{(i_q+1)} \qquad \text{for odd } n$$

$$q_{50}^a = (x_{i_q} + x_{(i_q+1)})/2 \qquad \text{for even } n$$

Where FIX() is the Fixum Function, which rounds downwards to the next Integer Value.

3.2.2 Analytical Statistics

In **Analytical Statistics** Relations within the data are uncovered. Here appear e.g. questions about the **Relations between two or more Variables** (*Correlation*, *Regression*), which go up to **Classification Methods** (*Minimum-Box-Classification*, *Distance-Classification*, *Maximum-Likelihood-Classification*).

3.3 Matrix Computation

In the age of computers Matrix Computation is the optimal form to represent and directly calculate with certain Equation Systems. The most Computer Algebra Systems are built on Matrices and their application, e.g. MATLAB, Maple and Mathematica (commercial) or FreeMat and Scilab (Open Source).

3.3.1 Vectors and Matrices

Matrices can be described as a rectangular arrangement of numbers and/or terms. These ordered numbers and/or terms are also called **Elements of the Matrix**.
In the following sections some fundamental Vector and Matrix properties are presented:

Order of a Matrix

The Order of a Matrix $\mathbf{A}$ is determined by the number of its Rows and Columns.
Hint: This term should not be confused with the Rank of a Matrix, which gives the linearly independent rows resp. columns.

Transpose of a Matrix

A Matrix $\mathbf{A}$ can be transposed by changing the arrangement of rows and columns. The transposed Matrix is written as $\mathbf{A'}$ or sometimes also as $\mathbf{A^T}$.

Vector as special case of a Matrix

A Matrix, which only has a single row or a single column, is called a Vector. There are **Row** and **Column Vectors**, dependent on which Dimension has the Order 1.

3.3.2 Matrix Algebra

For Matrices similar calculation rules can be defined as are existent for single numbers, which are also called **Scalars** concerning Matrices.

The Scalar Product (Dot Product)

A row vector $\mathbf{a'}$ and a column vector $\mathbf{b}$ of the same Order n build a Scalar Product as follows (the result is a number, a scalar):

$$\mathbf{a'} \cdot \mathbf{b} \quad = \quad a_1 \cdot b_1 + a_2 \cdot b_2 + \ldots + a_n \cdot b_n$$