Ausgewählte Kapitel der Parameterschätzung — Übung 5

Ausgabedatum: 22. June 2020 Abgabetermin: 6. July 2021, 18 Uhr

Maximum-a-posteriori-Schätzung und Bayessche Inferenz

The Gravity Recovery and Climate Experiment (GRACE) was a joint mission of NASA and the German Aerospace Center consisting of one pair of satellites aiming at measuring the (time-variable) gravitational field of the earth. GRACE was launched in March 2002 and ended its science mission in October 2017. The measurement principle was based on the precise estimation of the orbit and inter-range dynamics which were eventually related to the earth's gravitational field and mass distribution. In particular, GRACE has enabled scientists to monitor the mass change associated to the water cycle. Here, we will evaluate the trend of Total Water Storage Change (TWSC) over Greenland as derived from GRACE data. The TWSC signal in this region is dominated by the ice mass loss which results in a negative trend. The file `sim_TWS_Greenland.mat` available on ILIAS contains a **simulated** time series of TWSC over Greenland expressed in gigatons.

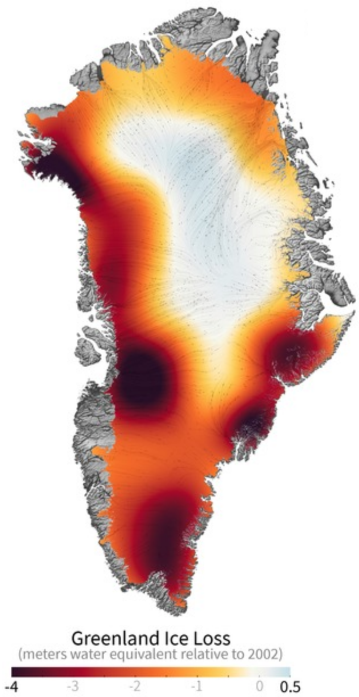

Figure 1: Map of the spatial distribution of ice mass loss trend (source *NASA*).

0.1 data gap filling

The objective of this section is to fill the data gaps occuring at months 46, 47, 115, 116, 117, 118 using a bayesian approach. For this, we will need a model of the time series and an estimation its parameters.

a) Plot the data. At first sight, the data shows essentially a trend and a yearly oscillation superposed on it. Hence the model M_1 we will use hereafter:

$$s(t) = at + b + A_o \cos(2\pi \frac{t}{T} + \varphi) + n(t)$$

where T is the period of 1 year, a, b, A_o and $\varphi \in \mathbb{R}$ are the unknown parameters and $n(t)$ is a centred, independent and normally distributed noise with an estimated standard deviation of 71 gigatons. Note that this model is non-linear in φ.

b) Prove that we can always write $A_o \cos(2\pi \frac{t}{T} + \varphi) = C \cos(2\pi \frac{t}{T}) + S \sin(2\pi \frac{t}{T})$ with $C, S \in \mathbb{R}$.

c) Determine the least-squares estimate of the parameters a, b, C and D. In particular, estimate the *a posteriori* variance-covariance matrix of the 4 parameters.

d) Now use the MAP approach to estimate $prob(d_i|data)$ $i \in \{1, 2, 3, 4, 5, 6\}$ where d_i is the i^{th} missing data. You can use as a prior distribution for each parameter a normal distribution with mean the value determined previously and as a variance the variance of the *a posteriori* variance.

e) How would you decrease the "weight" of the prior compared to likelihood function?

f) Finally, estimate $prob(d_i|data)$ $i \in \{1, 2, 3, 4, 5, 6\}$ where d_i is the i^{th} missing data.

0.2 Open (optional) question: is the ice sheet mass loss accelerating?

Previously, we modelled the multiannual evolution of the signal with a constant trend but one can argue that the trend is not exactly constant. Rather than imposing a constant trend, we would like to follow a data-driven approach that choose the appropriate model for the data. The competing model M_2 would be

$$s(t) = \alpha t^2 + \beta t + \gamma + C \cos(2\pi \frac{t}{T}) + S \sin(2\pi \frac{t}{T}) + n(t)$$

, which allows a constant acceleration.

1. Discuss how you would select the appropriate model.
2. One way to decide which one of the two models M_1 or M_2 is better, is to quantify the posterior ratio
$$\frac{prob(M_1|data)}{prob(M_2|data)}$$
If it is much greater than 1, then we will prefer the M_1 model. Using Bayes' theorem, marginalization and the product rule, try to evaluate the ratio.