Adaptive filtering in airborne gravimetry with hidden Markov chains.

Yury Bolotin ∗ Danila Doroshin ∗∗

∗ Mechanics and Mathematics Faculty, Moscow State University, 119991, Moscow, Russia (Tel: +7 (095) 9393383; e-mail: ybolotin@yandex.ru).
∗∗ Mechanics and Mathematics Faculty, Moscow State University, 119991, Moscow, Russia (Tel: +7 (095) 9393383; e-mail: doroshin@mail.ru).

Abstract: The paper studies filtering of airborne gravimetry data on a survey line. The aim is to build a filter that adapts to the nonhomogeneous gravity structure, which is induced by the unknown structure of Earth core, topography, etc. Gravity is modeled by a hidden Markov chain with finite number of states, each state corresponds to a certain type of gravity profile. Filtering is done in three steps: estimate the parameters of the Markov chain, determine moments of transitions between states, and construct the corresponding non-stationary Kalman filter. The algorithm takes into account the type of noise of the global positioning system. Tests on simulated and experimental data show that the filter diminishes over-smoothing and under-smoothing effects.

1. INTRODUCTION

The problem of airborne gravimetry (AG) is to determine the gravity anomaly (GA) on the trajectory of an aircraft using the observations of a gravimeter and GPS (Global Positioning System). Note that though the earth gravity is a function of position in space, in airborne gravity, for an aircraft moving along a straight survey line, it becomes a function of time. Therefore, the problem of AG is to determine GA as a function of time. When a gravity survey is done, these straight survey lines form a (usually) rectangular grid. With gravity estimates on the survey lines, the map of gravity in the survey region is calculated. However, this last problem is usually referred not to airborne gravimetry, but to geophysics and will not be discussed here.

The problem of AG can be reduced to filtering problem. This filtering is usually done with heuristic algorithms. See Alberts [2009], Hammada [1997]. It is also often done under stochastic assumptions, with a stochastic a priori model of GA, which is assumed to be a realization of a stationary random process. See Alberts [2009], Hammada [1997], Bolotin et al. [2005]. In this paper an adaptive model-based stochastic approach to AG filtering is suggested. Taking in mind non-homogeneous mass distribution in the Earth core, the gravity anomaly is assumed to be a priori a non-stationary random auto regression process, whose parameters are spatially varying and unknown.

This model of gravity is supplemented with a non-homogeneous in time model of the noise in the GPS data.

In more detail, the data are assumed to consist of stationary intervals, where the data are described by a mixture of moving averages (MA) with constant distributions parameters, and transient intervals, where the parameters of distributions change. The change in parameters is described by a Markov model. Thus the observed data are described by a hidden Markov model (HMM) based on a mixture of MA. Note that HMM in the classical sense describes processes, uncorrelated in time. Our model describes correlated processes. However, to avoid introducing new notation, we use the term HMM.

In the above settings the problem of airborne gravimetry can be reduced to adaptive filtering: evaluate the parameters of HMM, find the trajectory of HMM, and, finally, estimate the gravity anomaly. These three tasks are solved step-by-step with maximum likelihood approach (expectation-maximization algorithm (EM)), dynamic programming (Viterby algorithm), and optimal Kalman smoothing.

One problem in airborne gravimetry is very low signal-to-noise (SNR) ratio ($10^{-3} - 10^{-6}$), with most of the noise in the high frequency band. Thus to apply the adaptive filtering algorithm described above to the data, data regularization is required. Regularization is done with finite impulse response (FIR) filtering and resampling, such that the pay signal and the noise in the resulting signal are of the same order.

2. REDUCING THE AG PROBLEM TO ESTIMATION IN MIXTURE OF MA

A typical airborne gravimeter consists of a set of GPS receivers and a set of high precision gravity sensors (accelerometers) and angular velocity sensors (gyros) positioned on a gimbal platform. The accelerometers measure the specific force acting on the system proof mass (PM) projected to the instrument frame axes. By solving the aided (by GPS) inertial navigation problem, the specific
force is projected to the local geographical vertical. See Torge [1988]. The equation of motion of PM in projection on the local vertical can be written as follows (see Bolotin et al. [2005], Torge [1988]).

\[ V = f_3 + g_0 + \delta g + \Delta g_{ETV}. \]  \hspace{1cm} (1)

Here \( f_3 \) is the vertical component of the specific force acting on PM, \( V \) is the vertical component of velocity of PM, \( g_0 \) is the normal gravity force, \( \delta g \) is the gravity anomaly in free air (see Torge [1988]), \( \Delta g_{ETV} \) is the Eotvos correction term. See Torge [1988]. The terms \( g_0 \) and \( \Delta g_{ETV} \) can be computed with required accuracy with GPS readings, \( f_3 \) and \( V \) are measured by the instrument, while the anomaly \( \delta g \) is the unknown value which is to be estimated.

The observations in (1) are the estimate \( V' \) of vertical velocity of PM, computed using the GPS data, and the mean value of specific force \( f_3' \) during a sampling interval of the gravimeter. The sampling frequency of GPS is usually much lower than that of the gravimeter (see Bolotin et al. [2005], Torge [1988]), which permits to interpolate the gravimeter data to the GPS sampling rate, assuming that the above down sampling didn’t provoke aliasing (see Bolotin et al. [2005]). Choosing the value of \( m \), one can select an adequate model of gravity for geologically different areas.

The GPS noise will be modeled by a non-stationary Gaussian white noise in discrete time (see Bolotin et al. [2005]). This is in good correspondence with the physics of phase GPS observations (see Stepanov et al. [2002])

\[ E[\delta V(t)\delta V(s)] = \begin{cases} R(t), & t = s \\ 0, & t \neq s \end{cases} \]

Approximating the integral in (3) by \( \Delta t \delta g(t) \), applying \( \nabla^m \) to (3) and denoting \( x(t) = \nabla^m y(t) \), \( r(t) = \delta V(t) \), we get

\[ x(t) = \Delta t q(t) + \nabla^{m+1} r(t), \quad t = k\Delta t, \quad k = 0, 1, \ldots \]  \hspace{1cm} (6)

The left part of (6) is a function of observed quantities; the right part is a mixture of Gaussian moving averages. Thus the problem of AG is reduced to estimating a mixture of MA.

From the methodological point of view we prefer to reduce (6) to a more general case of the observed signal being a sum of two MA

\[ x(t) = x_r(t) + x_q(t), \]  \hspace{1cm} (7)

\[ x_q(t) = \sum_{l=0}^{K} c_l q(t-l\Delta t), \quad x_r(t) = \sum_{l=0}^{L} d_l r(t-l\Delta t). \]

The coefficients \( c_l, d_l \) will be referred to as the weights of the MA.

3. REGULARIZATION OF DATA

In airborne gravimetry the SNR is very low, as the GPS noise is 5-6 orders higher than the pay signal, the gravity anomaly. Thus the problem of evaluating the gravity is very ill-conditioned, which in our settings means that the second term in (6), (7) is many orders more than the first term. To regularize the problem, it is reduced to the lower frequency interval, where the energies of two signals are comparable. This is done by FIR smoothing and down sampling. Smoothing is done with the Kaiser window filter (see Hamming [1989]), whose frequency response \( H(\omega) \) is selected according to the a priori assumptions on the signal variances \( Q \) and \( R \) (at the regularization stage \( Q \) and \( R \) are assumed constant). The cutoff frequency of the filter \( \omega_c \) is selected for the energy of GA and GPS noise in (6), (7) in the frequency band \( |\omega| < \omega_c \), to be comparable. As a result, the MA weights in (7) are convoluted with the filter weights.

Next, the data are downsampled \( n \) times to reach the sampling interval \( \Delta t' = n\Delta t \) such that the Nyquist frequency \( \omega_N' = \pi/\Delta t' \) of the down sampled signal \( x'(t) \), \( t = k\Delta t', \quad k = 0, 1, \ldots \) is equal to the filter stopband frequency \( \omega_{stop} \) (see Hamming [1989]). Assuming that the above down sampling didn’t provoke aliasing (see)

\[ \nabla^m \delta g(t) = q(t), \]

\[ E[q(t)q(s)] = \begin{cases} Q(t), & t = s \\ 0, & t \neq s \end{cases} \]

where \( \nabla^m \) is the \( m \)-th difference, and \( q(t) \) is a Gaussian white noise in discrete time. Assumption (4) is shown to be adequate (see Bolotin et al. [2005]).

The left part of (6) is a function of observed quantities; the right part is a mixture of Gaussian moving averages. Thus the problem of AG is reduced to estimating a mixture of MA.

From the methodological point of view we prefer to reduce (6) to a more general case of the observed signal being a sum of two MA

\[ x(t) = x_r(t) + x_q(t), \]  \hspace{1cm} (7)

\[ x_q(t) = \sum_{l=0}^{K} c_l q(t-l\Delta t), \quad x_r(t) = \sum_{l=0}^{L} d_l r(t-l\Delta t). \]

The coefficients \( c_l, d_l \) will be referred to as the weights of the MA.

3. REGULARIZATION OF DATA

In airborne gravimetry the SNR is very low, as the GPS noise is 5-6 orders higher than the pay signal, the gravity anomaly. Thus the problem of evaluating the gravity is very ill-conditioned, which in our settings means that the second term in (6), (7) is many orders more than the first term. To regularize the problem, it is reduced to the lower frequency interval, where the energies of two signals are comparable. This is done by FIR smoothing and down sampling. Smoothing is done with the Kaiser window filter (see Hamming [1989]), whose frequency response \( H(\omega) \) is selected according to the a priori assumptions on the signal variances \( Q \) and \( R \) (at the regularization stage \( Q \) and \( R \) are assumed constant). The cutoff frequency of the filter \( \omega_c \) is selected for the energy of GA and GPS noise in (6), (7) in the frequency band \( |\omega| < \omega_c \), to be comparable. As a result, the MA weights in (7) are convoluted with the filter weights.
Hamming [1989]), the power spectral density (PSD) of the
down sampled signal $x'(t)$ can be written in the frequency
band $|\omega| < \omega_N$ as $S'_q(\omega) = S_q(\omega) + S'_r(\omega)$, where

$$S_q(\omega) = \frac{Q}{2\pi} |H(\omega)H_r(\omega)|^2, \quad S'_r(\omega) = \frac{R}{2\pi} |H(\omega)H_e(\omega)|^2.$$

Here $H_q(\omega)$, $H_r(\omega)$ are the transfer functions of MA, which
can be determined from the weights $c_l$, $d_l$ in (7) by the
formulas of the complex Fourier series. See Katayama [2005].

After filtering and down sampling the data loses the
structure of MA. To return the structure of MA, we
construct a new (similar to (7)) approximating MA model
for the downsampled data as

$$x'(t) = x'_q(t) + x'_r(t) + e'(t),$$

$$x'_q(t) = \sum_{i=0}^{K'} c'_q(t - l\Delta' t), \quad x'_r(t) = \sum_{i=0}^{L'} d'_r(t - l\Delta' t).$$

Here $q'(t)$, $r'(t)$ are the generating white noises in the
downsampled space, $e'(t)$ is the approximation error, which
is assumed to be uncorrelated with $q'(t)$ and $r'(t)$.

Approximation is done with spectral technique. PSD of the
signal $x'(t)$, presented in the MA form (9), can be written
in the frequency band $|\omega| < \omega_N$ as the sum of power
spectral densities of its components $S'_q(\omega) = S_q(\omega) +
S'_r(\omega)$, where $S_q(\omega)$ is PSD of the error $e'(t)$,

$$S'_q(\omega) = \frac{Q}{2\pi} |H_q(\omega)|^2, \quad S'_r(\omega) = \frac{R}{2\pi} |H_r(\omega)|^2,$$

$H_q(\omega)$, $H_r(\omega)$ are the transfer functions of MA which
can be determined from the weights $c'_m$, $d'_m$ by the formulas
for the complex Fourier series. See Katayama [2005].

The orders $K'$, $L'$ of MA and their weights $c'_l$, $d'_l$
are chosen for the power spectral densities $S'_q(\omega)$, $S'_r(\omega)$
of the components in mixture (9) to approximate the components
of PSD of the down sampled signal $x'(t)$ in (8).

$$\int_{-\omega_N}^{\omega_N} |H_q(\omega)H(\omega) - H'_q(\omega)|^2 d\omega \rightarrow \min,$$

$$\int_{-\omega_N}^{\omega_N} |H_r(\omega)H(\omega) - H'_r(\omega)|^2 d\omega \rightarrow \min.$$
explicitly. Let us denote $X_{t_1}^{t_2} = [x(t_1), ..., x(t_2)]$, $S_{t_1}^{t_2} = [s(t_1), ..., s(t_2)]$. Let us denote the correlation radius of the process $x(t)$ in (10) as $p = \max\{K, L\}$. Note that $x(t)$ correlates with $2p + 1$ observations $X_{t-p}^{t+p}$. Note also that distribution of the process $x(t)$ is determined by statistical characteristics of the gain noises at moments $[t-p, ..., t]$, or, equivalently, by the set of states $S_{t-p}^t$. Let $\gamma_t(S_{t-p}^t) = P_S[S_{t-p}^t | X, \Theta_m]$, be conditional probability of the state of the system with full set of observations $X$ and parameters of the model $\Theta_m$. With the above notation (11) can be written as

$$U(\Theta, \Theta_m) = \sum_{s(-p)} \log \pi_s(-p) \gamma_p(s(-p)) +$$

$$+ \sum_{t=-p+1}^{T-1} \log a_{s(t-1), s(t)} \gamma_t(s(t-1), s(t)) +$$

$$+ \sum_{t=0}^{T-1} \sum_{S_{t-2p}} \log f_{X}(x(t) | X_{t-p}^{t-1}, S_{t-2p}^t, \Theta) \gamma_t(S_{t-2p}^t)$$

We see that (13) is a sum of two components depending on two different groups of parameters. Thus optimization in these two groups can be done independently. To estimate the transition matrix and the initial Markov chain probabilities the EM-algorithm is used. During optimization it is necessary to take into account restrictions on the parameters: the sum of all elements in any transition matrix row equals and initial probabilities are: $\sum_i a_{ij} = 1$, $\sum_i \pi_i = 1$. Using the Lagrange multipliers and subsequent unconditional optimization, the optimal chain parameters for the $m + 1$-th iteration of (12) can be written as

$$a_{i,j}^{m+1} = \frac{\sum_{t=-p+1}^{T-1} \gamma_t(i, j)}{\sum_{i=-p}^{T-2} \gamma_t(i)}$$

$$\pi^{m+1} = \gamma_p(i)$$

Here $\gamma_p(i)$ is the conditional probability calculated for the initial time moment, which is $t = -p$, since the first observation $x(0)$ in (10) is determined by the system noise at $p$ previous moments. Note also that the formulas (14) do not differ from similar formulas for the standard HMM parameter re-estimation. See Vaseghi [2006].

Optimization of MA parameters with the algorithm brings up a nonlinear optimization problem. This problem is solved by numeric optimization of (13) with modified gradient coordinate descent. See Bonnas et al. [2003]. Initial conditions were set to $\Theta_m$. Iteration stopped when the relative change of was below the given threshold.

All formulas (13), (14) for re-estimation are based on calculating conditional probabilities $\gamma_t(S_{t-p}^t)$ for all moments $t$ and all possible state sequences $S_{t-p}^t$. These probabilities are calculated by the forward-backward algorithm (see Vaseghi [2006]) modified for the case of HMM built on a MA mixture. The probability $\gamma_t$ may be presented in the following way:

$$\gamma_t(S_{t-2p}^t) = \frac{f_{X,S}(S_{t-2p-1}^t, X(\Theta_m))}{f_X(X(\Theta_m))}$$

The main idea used to calculate $\gamma_t$ in the forward-backward algorithm is that the numerator of (15) may be factorized into a combination of terms dependent only on the observations made before moment $t$ and ones dependent only on the observations made after this moment. To do this the so-called forward and backward probabilities are used. The forward probabilities depending on the observations before $t$ are defined by the following expression:

$$\alpha_t(S_{t-2p+1}^t) = f_X, S(X_{t-p}^t, S_{t-2p+1}^t | \Theta_m)$$

The backward probabilities depending on the observations after $t + r$ are defined by formula

$$\beta_t(S_{t-2p+1}^t) = f_X, S(X_{t+p}, S_{t-2p+1}^t | s(t), \Theta_m)$$

The forward probabilities at the moment $t$ may be calculated iteratively from the forward probabilities at the previous moment as

$$\alpha_t(S_{t-2p+1}^t) = \sum_{t=-p+1}^T f_X(x(t) | X_{t-p}^{t-1}, S_{t-2p}^t, \Theta) \alpha_{t-1}(S_{t-2p-1}^t) \beta_{t-1}(S_{t-2p+1}^t)$$

The initial values for the forward probabilities may be found from the following equation:

$$\alpha_{-p-1}(S_{-p}^{-1}) = f_X(X_0^{-1}, S_{-p}^{-1}, \Theta_m)$$

Using forward probabilities calculated for the finite moment of time we get the likelihood function in (15) as

$$f_{X}(X | \Theta_m) = \sum_{S_{T-2p}^T} \alpha_{T-1}(S_{T-2p}^T)$$

The backward probabilities are calculated similarly. Iterations go on in the reversed time, from the moment $t + 1$ to the moment $t$ as

$$\beta_t(S_{t+2p}^{t+1}) = \sum_{t+2p} f_X(x(t+p) | X_{t+2p}^{t+1}, S_{t+2p}^{t+1}, \Theta_m) \times a_{s(t), s(t+1)} \beta_{t+1}(S_{t+2p}^{t+1})$$

The initial values of the backward probabilities are determined by the following formula:

$$\beta_{T-2p}(S_{T-2p}^{T-1}) = f_X(X_{T-2p}^{T-1}, S_{T-2p}^{T-1}, \Theta_m) \times \prod_{k=T-2p+1}^{T-1} a_{s(k-1), s(k)}$$

Finally, the conditional probabilities (15) can be written using the forward and backward probabilities (16), (17) as

$$\gamma_t(S_{t-2p}^t) =$$

$$= \left( \sum_{t+2p} f_X(X_{t+2p-1}^t, X_{t+2p}, S_{t+2p}^{t+1}, \Theta_m) \right) \times$$

$$\times a_{s(t-1), s(t)} \beta_{t-1}(S_{t-2p+1}^{t-2p-1}) / f_X(X(\Theta_m))$$

Here the term

$$f_X(X_{t+2p-1}^t, X_{t+2p}, S_{t+2p}^{t+1}, \Theta_m)$$

represents the conditional probability density of data block $X_{t+2p-1}^t$ with known previous and foregoing data blocks.
The state sequence derived in the recognition process as \( \tilde{S} \) and the set of parameters derived in the training algorithm as \( \Theta \). Then the solution of (30) can be obtained by the well-known formula \( w_t = R_{XX}^{t} r_x(t)X \). See Vaseghi [2006].

Modulo a constant shift, the estimate \( \widetilde{\delta g}(t) \) of GA can be obtained from \( \tilde{x}_q(t) \) by simple integration. To estimate the constant shift and to do additional regression analysis of the data, a Kalman smoother can be used. See Bolotin et al. [2005].
5. RESULTS AND CONCLUSIONS

Let us present the results of processing the data provided by the GT1A gravimeter (see Berzhitsky et al. [2002]) on a straight survey line. The aircraft was moving at the speed of 100 m/s. The Ashtech GPS receivers frequency was 10 Hz, the gravimeter sampling frequency was 18.24 Hz. For GPS operation in differential mode, two base stations were used. The data were given to the authors by the company Aerogeofizika Ltd (Russia).

GPS data processing, aided inertial navigation, and calculation of inertial corrections was done with GTNAV - GTGRAV software. See Berzhitsky et al. [2002]. The GPS velocity noise RMS of the raw data was in the range $1 \cdot 10^{-2} - 2 \cdot 10^{-4} \text{m/sec}$. To regularize the data it was smoothed with 1 km spatial resolution. The GA was described by the first integral of a white noise whose intensity was given by the Markov model with two states. The GPS noise was supposed to be stationary. The moving averages of the seventh order were used to describe the regularized anomaly and the GPS noise. The training algorithm determined both the gravity anomaly variances and the GPS noise variance as $Q_1 = 4.1417 \cdot 10^{-12} \text{m}^2/\text{sec}^2$, $Q_2 = 3.5330 \cdot 10^{-11} \text{m}^2/\text{sec}^2$, $R = 1.6788 \cdot 10^{-4} \text{m}^2/\text{sec}^2$. Intervals with different anomaly intensity are shown in Fig. 1.

Let us compare the obtained GA with the one given by a stationary filtering algorithm. See Bolotin et al. [2005]. The GA model parameters for the stationary algorithm were taken equal to $Q_2$, $R$. Fig. 2 shows fragments of GA estimates zooming in the rectangle in Fig. 1. This fragment corresponds to the lower value of gravity intensity. It may be seen from Fig. 2 that the inaccuracy of the chosen parameters for the stationary algorithm led to undersmoothing. See Bolotin et al. [2005]. The error in evaluating GA was estimated as RMS of the difference of GA on intersecting flight lines. The adaptive algorithm was found out to perform similar to the stationary algorithm on the intervals of high intensity of gravity, and to perform 50% better on the intervals of low intensity of gravity.

REFERENCES


