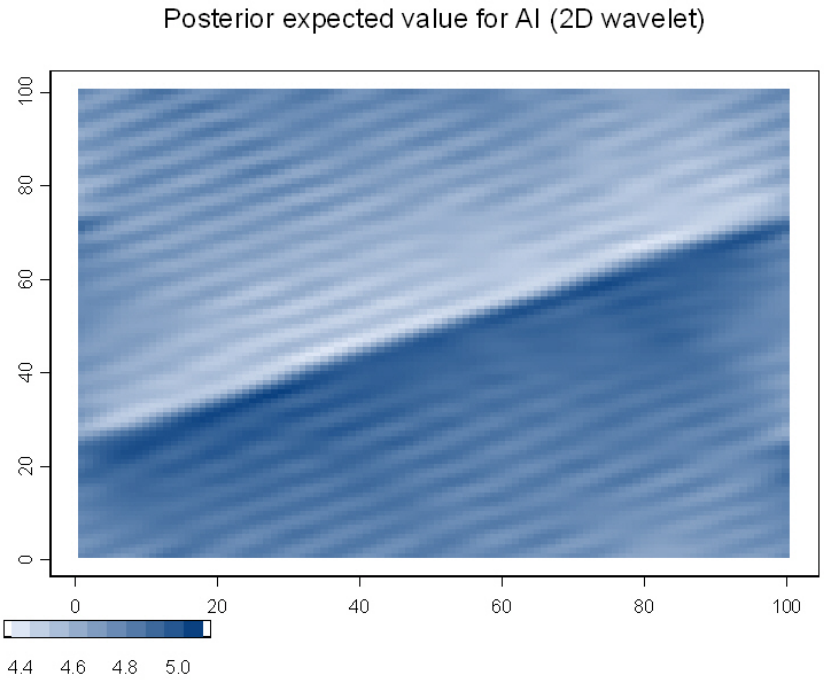
Norsk Regnesentral
NORWEGIAN COMPUTING CENTER

Note

Simple SIMPLI Inversion



Note no SAND/08/08
Authors Frode Georgsen
Odd Kolbjørnsen

Date 2nd October 2008

Norwegian Computing Center

Norsk Regnesentral (Norwegian Computing Center, NR) is a private, independent, non-profit foundation established in 1952. NR carries out contract research and development projects in the areas of information and communication technology and applied statistical modeling. The clients are a broad range of industrial, commercial and public service organizations in the national as well as the international market. Our scientific and technical capabilities are further developed in co-operation with The Research Council of Norway and key customers. The results of our projects may take the form of reports, software, prototypes, and short courses. A proof of the confidence and appreciation our clients have for us is given by the fact that most of our new contracts are signed with previous customers.

Norsar

NORSAR is an independent research foundation specializing in commercial software solutions and research activities within applied geophysics and seismology. NORSAR operates some of the world's largest seismological observatories, and has more than 35 years of experience in developing and supporting advanced seismological data processing and analysis schemes. NORSAR is designated as the Norwegian National Data Center (NDC) for verification of compliance with the Comprehensive Nuclear-Test-Ban treaty (CTBT). NORSAR has since 1977 been engaged in research and development in the fields of seismic prospecting, and today offer state-of-the-art 2D and 3D seismic modelling software worldwide.

Norsk Regnesentral
Norwegian Computing Center
Postboks 114, Blindern
NO-0314 Oslo, Norway

Besøksadresse
Office address
Gaustadalléen 23
NO-0373 Oslo, Norway

Telefon · telephone
(+47) 22 85 25 00
Telefaks · telefax
(+47) 22 69 76 60

Internett · internet
www.nr.no
E-post · e-mail
nr@nr.no

Title	**Simple SIMPLI Inversion**
Authors	**Frode Georgsen** <frode.georgsen@nr.no> **Odd Kolbjørnsen**
Date	2nd October 2008
Publication number	SAND/08/08

Abstract

A simplified Bayesian inversion with a spatial wavelet is presented in theory and with test examples

Keywords	Bayesian inversion, Spatial wavelet, Fast Fourier Transform
Target group	Users of SeisRox, Partners in the BIP project
Availability	Open
Project	Fast Elastic Inversion of Multi-offset Prestack Depth Migrated Seismic Data
Project number	492001
Research field	Stochastic seismic inversion
Number of pages	22

Contents

1 Introduction

The work presented in this note describes preliminary testing performed in the early stage of the research project defined by the BIP ("Brukerstyrt Innovasjons Prosjekt") agreement between Norsk Forskningsråd and Norsar Innovation AS. Norwegian Computing Center (NR) is one of the research partners in this project. The main objective in the project is to develop a method for elastic inversion that combine the very fast inversion technique (CRAVA) developed by NR with a novel technique for modelling prestack depth migrated seismic data (SIMPLI) developed by the other research partner NORSAR. The theory behind the Bayesian Inversion used in CRAVA is described in Buland et al. (2003) and a presentation of a real case where the program has been used is given in Dahle et al. (2007). The SIMPLI technology is described in Lecomte (2008) showing how the use of a spatial filter based on the concept of point-spread functions (PSF) as an alternative to traditional 1D convolution in the forward modelling, gives more realistic simulated PSDM cubes. The main idea in the BIP-project is to perform the bayesian inversion on PSDM cubes by using PSFs, or the time-domain analog, 3D wavelets, and to incorporate this method in the existing seismic analysis software (SeisRox).

As a simple test of concept the Bayesian Inversion has been performed on a 2D target area using both a 1D- and a 2D-wavelet. This note describes the application of formulas given in Buland et al. (2003) and how this theory easily can be extended to a wavelet of higher dimension. Some test examples are presented. In addition an inversion with a bayesian trend in the elastic parameter is presented.

In the appendices some theory on Fourier transformation and matrix operations is given. This theory is either used in the present paper or in Buland et al. (2003).

2 Seismic convolutional model

The Fourier transform of the convolutional model for the seismic data is given by expression (9) in Buland et al. (2003). When $\mathbf{k}$ is the spatial frequency vector, ω the temporal frequency and θ the reflection angle, this expression can alternatively be written as

$$\tilde{d}(\mathbf{k},\omega,\theta) = \tilde{s}(\mathbf{k},\omega,\theta)\mathbf{a}(\theta)\tilde{\mathbf{m}}'(\mathbf{k},\omega) + \tilde{e}(\mathbf{k},\omega,\theta) \tag{1}$$

where $\tilde{}$ refers to the Fourier domain, d the seismic data, s the wavelet, $\mathbf{a}$ the reflection coefficient based on the weak contrast approximation (Aki and Richards, 1980) and $\mathbf{m}$ the natural logarithm of the elastic parameters. Note that, unlike in the 1D convolution case, the wavelet, s, will generally also be dependent on the spatial frequency k. When operating on a discrete grid the differentation of $\mathbf{m}$ w.r.t t becomes a difference,

$$\frac{\partial}{\partial t} m(\mathbf{x},t) = m(\mathbf{x},t+1) - m(\mathbf{x},t)$$

Taking the DFT of this gives (when leaving out $\mathbf{x}$ and $\mathbf{k}$)

$$
\begin{aligned}
DFT(\frac{\partial}{\partial t}m(t)) &= \tilde{m}'(\omega) = \sum_{t=0}^{N-1}(m(t+1)-m(t))\exp\{-2\pi i\frac{\omega t}{N}\} \\
&= \sum_{t=0}^{N-1} m(t+1)\exp\{-2\pi i\frac{\omega t}{N}\} - \sum_{t=0}^{N-1} m(t)\exp\{-2\pi i\frac{\omega t}{N}\} \\
&= \sum_{s=1}^{N} m(s)\exp\{-2\pi i\frac{\omega(s-1)}{N}\} - \tilde{m}(\omega) \\
&= \sum_{s=0}^{N-1} m(s)\exp\{-2\pi i\frac{\omega s}{N}\}\exp\{2\pi i\frac{\omega}{N}\} - \tilde{m}(\omega) \\
&= \tilde{m}(\omega)(\exp\{2\pi i\frac{\omega}{N}\} - 1) \\
&= \tilde{m}(\omega)(\exp\{-2\pi i\frac{\omega}{N}(-1)\} + (-1)\exp\{-2\pi i\frac{\omega}{N}(0)\} \\
&= \tilde{m}(\omega)\sum_{j=0}^{N-1} h(j)\exp\{-2\pi i\frac{\omega}{N}(j)\} \\
&= \tilde{m}(\omega)\tilde{h}(\omega),
\end{aligned}
$$

where

$$
\begin{aligned}
h(0) &= -1 \\
h(-1) &= h(N-1) = 1 \\
h(j) &= 0 \text{ for } j \in \{1,\ldots,N-2\}.
\end{aligned}
$$

Then the convolutional model from expression (1) can be written as

$$\tilde{d}(\mathbf{k},\omega,\theta) = \tilde{h}(\mathbf{k},\omega)\tilde{s}(\mathbf{k},\omega,\theta)a(\theta)\tilde{m}(\mathbf{k},\omega) + \tilde{e}(\mathbf{k},\omega,\theta)$$

For the cases treated in the remaining we assume $\theta = 0$, so $a(\theta) = \frac{1}{2}$. Setting $g(\mathbf{k},\omega) = \frac{1}{2}\tilde{h}(\mathbf{k},\omega)\tilde{s}(\mathbf{k},\omega)$ we then get

$$\tilde{d}(\mathbf{k},\omega) = g(\mathbf{k},\omega)\tilde{m}(\mathbf{k},\omega) + \tilde{e}(\mathbf{k},\omega) \tag{2}$$

3 Bayesian inversion with 2D wavelet in 2D target area

In the following the theory from above will be applied in a 2D space (i.e. x respectively k is one-dimensional) with a 2D wavelet, restricted to one elastic parameter,

$$m(x,t) = \log(Z_P(x,t)). \tag{3}$$

where Z_p is the acoustic impedance.

3.1 Spatial wavelet

The 2D wavelet used below is defined as

$$s(x,t) = \exp\{-(\frac{x}{W_x})^2\}[1 - 2\pi^2\omega_0^2 t^2]\exp\{-\pi^2\omega_0^2 t^2\}, \tag{4}$$

with x measured i meters, t in seconds, W_x being the wavelet range in x and ω_0 the peak frequency. The vertical part of the wavelet is a Ricker. Figure 1 shows the vertical, horizontal and combined wavelets.

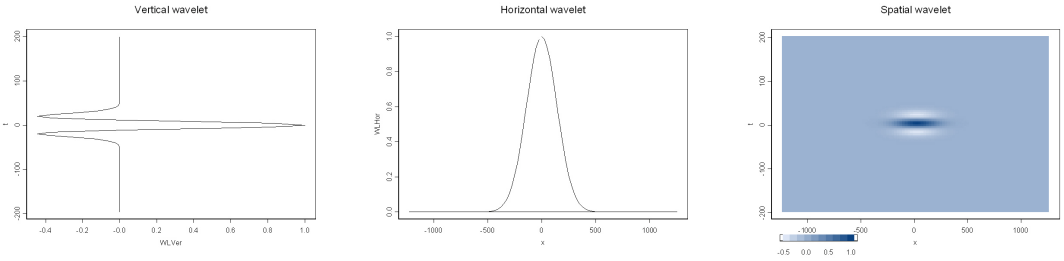


Figure 1. Spatial wavelet used in test examples. Left: Vertical wavelet. Centre: Horizontal wavelet. Right: Compined spatial (2D) wavelet.

3.2 Prior parameter model and seismic model

Let $\mathbf{m} = \{m(x,t);\ \forall(x,t)\}$ be the vector of the elements in expression (3) defined on a regular grid with $n = n_x n_t$ nodes where n_x and n_t are the number of gridnodes and Δx and Δt the sampling distances in the x- and t-directions respectively. To use the theory of circulant covariance matrices from appendix B.5 , $\mathbf{m}$ is assumed to be sampled on a circle in both x and t, meaning that the first and the last node in the sample in each direction are neighbours.

The model parameter $\mathbf{m}$ is assumed being Gaussian with a stationary and homogenous covariance function

$$\mathbf{m} \sim N_n(\boldsymbol{\mu}_m, \sigma_m^2 \boldsymbol{\Upsilon}_m), \tag{5}$$

with $\boldsymbol{\mu}_m = \{\mu(x,t);\ \forall(x,t)\}$ being the expectation vector for $\mathbf{m}$, σ_m^2 being the variance for all $m(x,t)$ and $\boldsymbol{\Upsilon}_m = \{\nu_m(|x-x'|, |t-t'|),\ \forall(x,t),(x',t')\}$ being the correlation matrix for $\mathbf{m}$. To make sure that this is a proper circulant covariance matrix the ranges must be less than $\Delta x \cdot n_x/2$ and $\Delta t \cdot n_t/2$. We then have, for the Fourier transformed parameter vector,

$$\tilde{\mathbf{m}} \sim N_n(\tilde{\boldsymbol{\mu}}_m, \sigma_m^2 \tilde{\boldsymbol{\Lambda}}_m) \tag{6}$$

with $\tilde{\boldsymbol{\mu}}_m$ being the Fourier transformed expectation vector and$\tilde{\boldsymbol{\Lambda}}_m = \{\tilde{\lambda}_m(k,\omega);\ \forall(k,\omega)\}$ being the diagonal matrix with the eigenvalues of $\boldsymbol{\Upsilon}_m$ multiplied by n on the diagonal in accordance with expression (B.5). The eigenvalues are found by taking the DFT of the first row as shown in appendix B.4. Expression (6) implies that each of the frequency components of $\tilde{\mathbf{m}}$ are independent

$$\tilde{m}(k,\omega) \sim N(\tilde{\mu}_m(k,\omega), \sigma_m^2 \tilde{\lambda}_m(k,\omega)).$$

By the same argument we get that the Fourier transformed error terms in the convolutional model (2) are componentwise independent

$$\tilde{e}(k,\omega) \sim N(0, \sigma_e^2 \tilde{\lambda}_e(k,\omega)),$$

where σ_e^2 is the noise variance and $\{\tilde{\lambda}_e(k,\omega);\ \forall(k,\omega)\}$ are the eigenvalues of the correlation matrix for the noise term $\mathbf{e} = \{e(x,t)\ \ \forall(x,t)\}$ multiplied by n.

This gives that the Fourier transformed seismic data defined in expression (2) are componentwise independent

$$\tilde{d}(k,\omega) \sim N(\tilde{\mu}_d(k,\omega), \tilde{\sigma}_d^2(k,\omega))$$

where

$$\begin{aligned} \tilde{\mu}_d(k,\omega) &= g(k,\omega)\tilde{\mu}_m(k,\omega) && (7)\\ \tilde{\sigma}_d^2(k,\omega) &= g(k,\omega)\sigma_m^2\tilde{\lambda}_m(k,\omega)g(k,\omega)^* + \sigma_\epsilon^2\tilde{\lambda}_\epsilon(k,\omega) && (8) \end{aligned}$$

and the crosscovariance between the seismic data and the model parameter is

$$\mathrm{Cov}[\tilde{d}(k,\omega), \tilde{m}(k,\omega)] = \tilde{\sigma}_{d,m} = g(k,\omega)\sigma_m^2\tilde{\lambda}_m(k,\omega). \quad (9)$$

3.3 Posterior model

The Fourier transformed observed seismic data in position (k,ω) is denoted $\tilde{d}_{obs}(k,\omega)$. Using formulas for conditional multinormal distribution (see for instance Johnson and Wichern (1988)) and expressions (7 - 9) gives the following componentwise posterior model

$$\tilde{m}(k,\omega)|\tilde{d}_{obs}(k,\omega) \sim N(\tilde{\mu}_{m|d}(k,\omega), \tilde{\sigma}_{m|d}^2(k,\omega))$$

where

$$\begin{aligned} \tilde{\mu}_{m|d}(k,\omega) &= \tilde{\mu}_m(k,\omega) + (\tilde{\sigma}_{d,m})^* \frac{1}{\tilde{\sigma}_d^2(k,\omega)}(\tilde{d}_{obs}(k,\omega) - \tilde{\mu}_d(k,\omega))\\ \tilde{\sigma}_{m|d}^2(k,\omega) &= \sigma_m^2\tilde{\lambda}_m(k,\omega) - (\tilde{\sigma}_{d,m})^* \frac{1}{\tilde{\sigma}_d^2(k,\omega)}(\tilde{\sigma}_{d,m}). \end{aligned}$$

4 Examples

The inversion with spatial wavelet is performed on three different examples. Two sets of synthetic seismic data are produced by performing a forward modelling using the convolutional model with a spatial (2D) wavelet. In the three examples the inversion method described in the previous section is performed and the posterior expectation $\tilde{\mu}_{m|d}$ is calculated and used as a prediction for the elastic parameter. In each example an inversion with the traditional 1D wavelet is also performed, and the result is used for comparison.

4.1 Forward modelling

Synthetic seismic data are constructed by the convolutional model (2) on a grid with $n_x = n_t = 100$, and grid distance 25 meters in x and 0.004 seconds in t. The wavelet used

in the forward modelling is the one described in expression (4) with $W_x = 200$ meters and $\omega_0 = 20$ hz. The error term is white noise with zero expectation and standard deviation of 0.01. In both datasets the target area is divided in two parts M_1 and M_2 with a low and a high constant value for $m(x,t)$ in each part, defined by

$$m(x,t) = \begin{cases} \ln 5.0 & \text{if } (x,t) \in M_1 \\ \ln 4.5 & \text{if } (x,t) \in M_2 \end{cases}$$

This gives one single reflector in the target area. In the first dataset the division follows a horizontal line between $t = 50$ and $t = 51$, and in the other the division line is diagonal starting between $t = 25$ and $t = 26$ on the left side ($x = 1$) of the target area and ending between $t = 75$ and $t = 76$ on the right side ($x = 100$). This is illustrated in Figure 2.

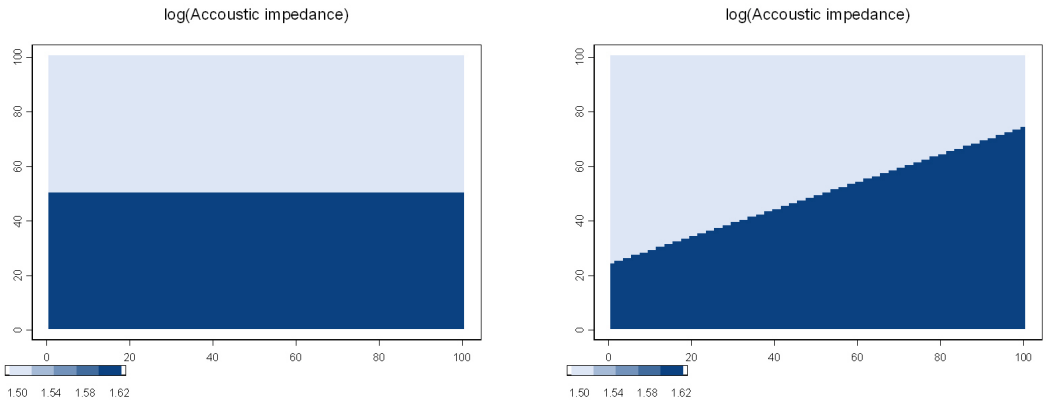


Figure 2. Model used for reflectors in constructing seismic datasets used for inversion.

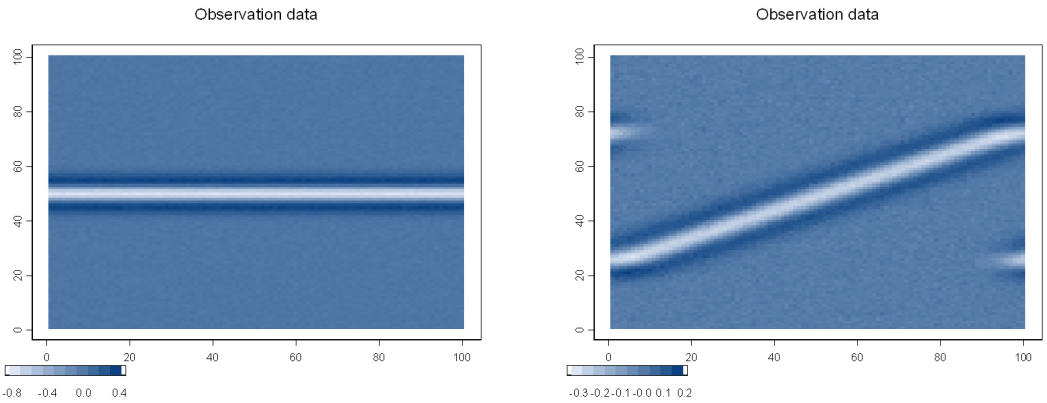


Figure 3. The synthetic seismic data corresponding to the reflectors in Figure 2.

The described procedure gives the synthetic seismic data shown in Figure 3. The signals visible on both sides of the figure to the right in Figure 3 are solely border effects caused by the fact that no padding has been used in the FFT-calculations.

4.2 Inversion

Three inversion example are constructed. In the first the inversion is performed on the dataset with a horizontal reflector. In the second example the dataset with diagonal re-

flector is used and with a correlation structure for the model parameter that does not follow the reflector. In the last example the dataset with diagonal reflector is used, but with a correlation function that is rotated to follow the reflector. In all examples the parameters for $\mathbf{m}$ in expression 5 are $\mu_m = 1.557$, $\sigma_m = 0.0527$. In case the correlation function is not rotated let Δx be the distance in x and Δt the distance in t between two points (x_1, t_1) and (x_2, t_2), so $\Delta x = |x_1 - x_2|$ and $|\Delta t = |t_1 - t_2|$. In the rotated case the slope of the rotated reflector in the (x, t)-system is $(50 \cdot 0.004)/2500 = 0.08$. Transformation gives $\Delta x = \cos[\arctan(0.08)] \cdot |x_1 - x_2| + \sin[\arctan(0.08)] \cdot |t_1 - t_2|$ and $dt = -\sin[\arctan(0.08)] \cdot |x_1 - x_2| + \cos[\arctan(0.08)] \cdot |t_1 - t_2|$. The correlation term in Υ_m in all examples is given by

$$\nu_m(\Delta x, \Delta t) = \exp\{-3.0 \cdot \sqrt{(\frac{\Delta x}{R_x})^2 + (\frac{\Delta t}{R_t})^2}\} \tag{10}$$

where $R_x = 1000$m and $R_t = 0.01$s.

For each of the three examples an inversion of the synthetic seismic data is performed. The horizontal reflection case is illustrated in Figures 4 and 5.

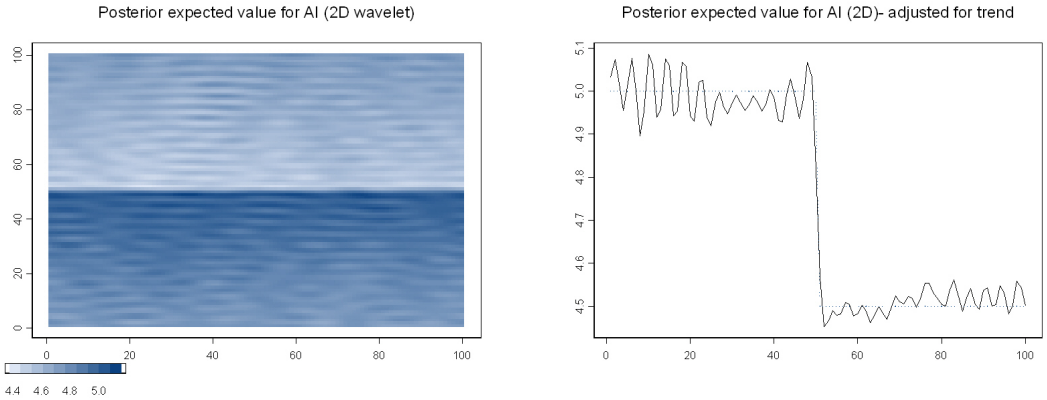


Figure 4. Results of the inversion with 2D wavelet for horizontal reflector. Left: Predicted value for accoustic impedance in the whole target. Right: Predicted value along one trace.

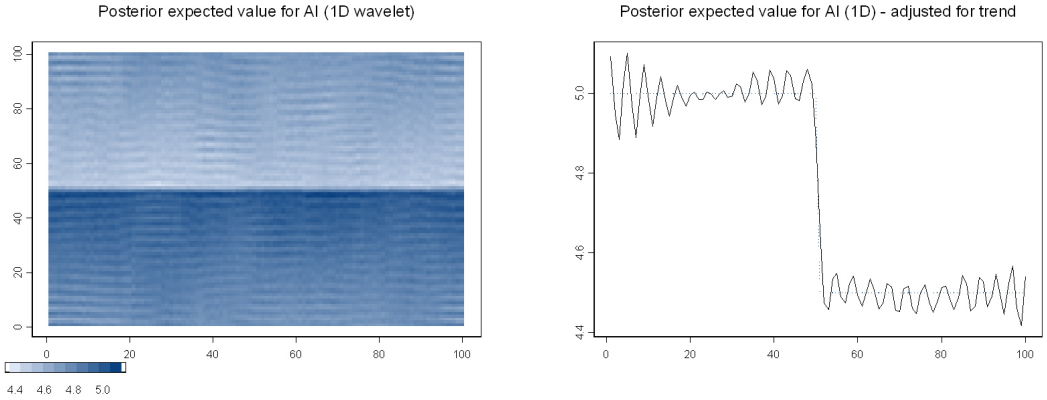


Figure 5. Results of the inversion with 1D wavelet for horizontal reflector . Left: Predicted value for accoustic impedance in the whole target. Right: Predicted value along one trace.

Figure 4 (left) shows the predicted values for the acoustic impedance in the whole target area after inversion with 2D wavelet. To the right is shown one single vertical trace after the inversion. In Figure 5 the corresponding figures are shown for the inversion with 1D wavelet. As one would expect, there is no significant difference between the two cases. When the reflector is horizontal there is no gain in using a 2D wavelet in the inversion. The inversion results from the case with dipped reflector, but non-rotated correlation function are shown in Figures 6 and 7.

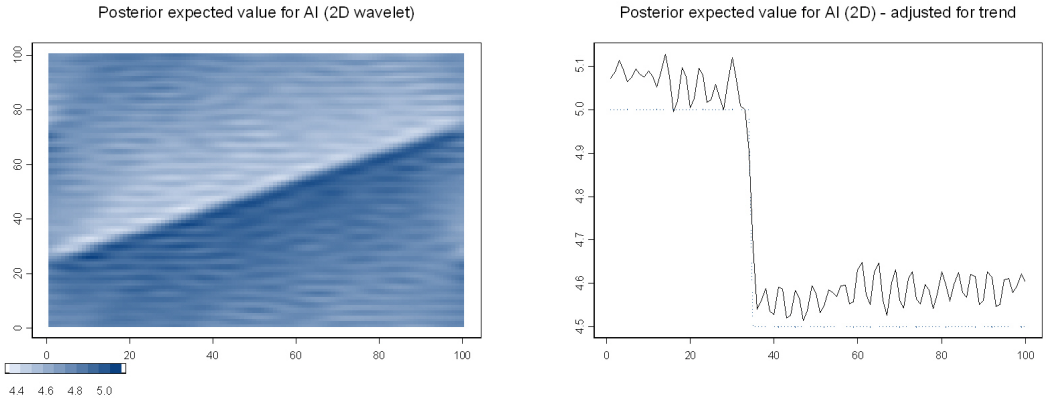


Figure 6. Results of the inversion with 2D wavelet, dipped refelctor and non-rotated correlation function. Left: Predicted value for accoustic impedance in the whole target. Right: Predicted value along one trace.

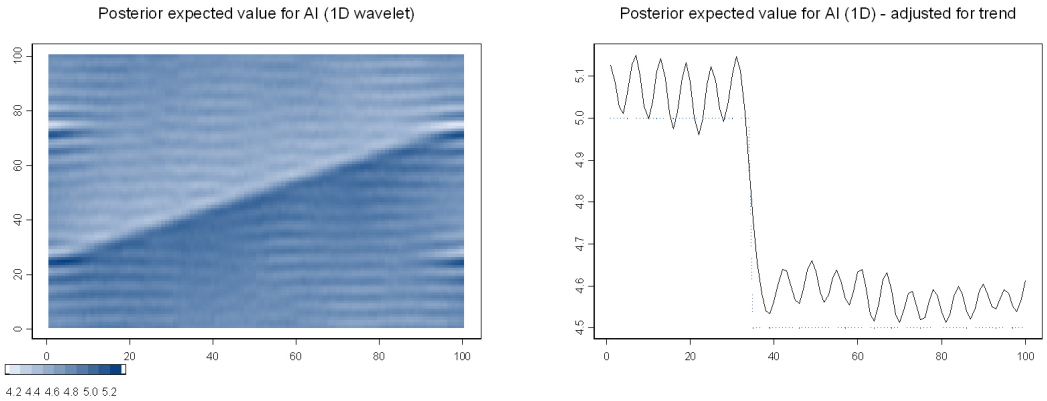


Figure 7. Results of the inversion with 1D wavelet, dipped reflector and non-rotated correlation function. Left: Predicted value for accoustic impedance in the whole target. Right: Predicted value along one trace.

Even if the differences between the inversions with 2D and 1D wavelet are not big, the figures indicates clearly that the inversion with 2D wavelet gives a sharper inversion result. The inversion results from the case with dipped reflector and rotated correlation function are shown in Figures 8 and 9.

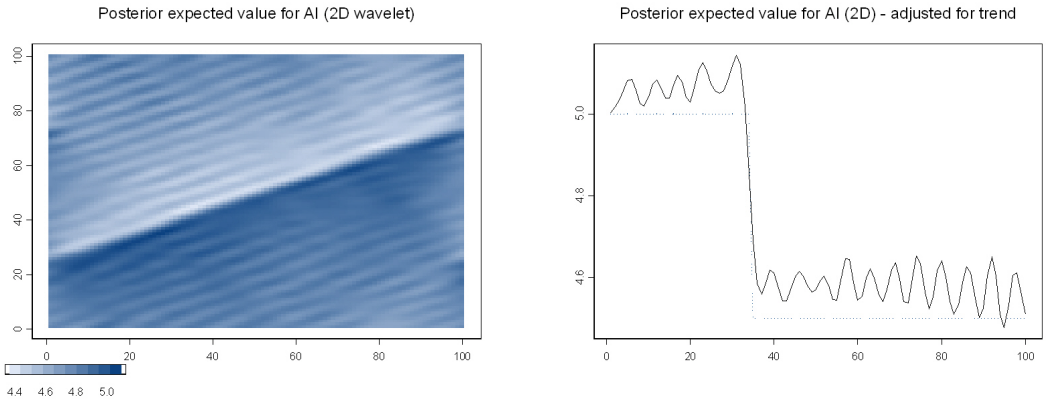


Figure 8. Results of the inversion with 2D wavelet, dipped reflector and rotated correlation function. Left: Predicted value for accoustic impedance in the whole target. Right: Predicted value along one trace.

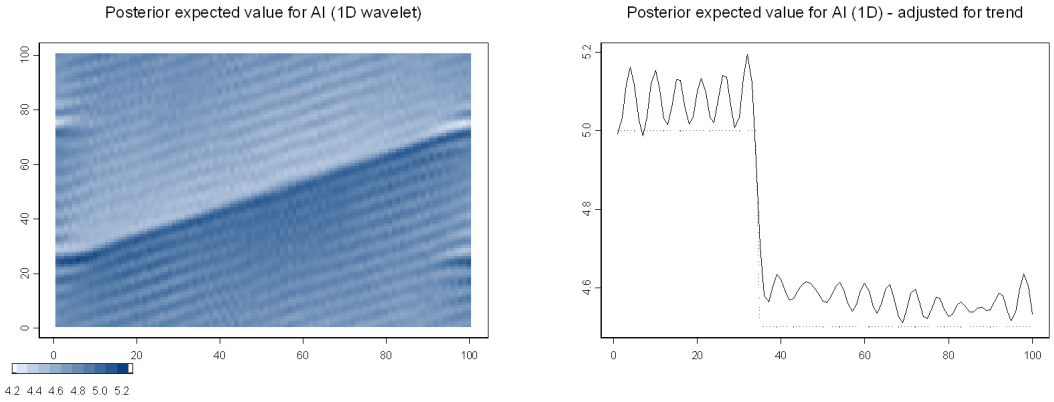


Figure 9. Results of the inversion with 1D wavelet, dipped reflector and rotated correlation function. Left: Predicted value for accoustic impedance in the wholde target. Right: Predicted value along one trace.

5 Bayesian trend in elastic parameter

Let $m(x,t)$ be the logarithm of the acoustic impedance defined by (3). A linear trend in $\mathbf{m} = \{m(x,t);\ \forall(x,t)\}$ is introduced by setting

$$m(x,t) = \sum_{l=1}^{L} f_l(x,t)\beta_l + r(x,t) = \mathbf{f}(x,t)\boldsymbol{\beta} + r(x,t), \tag{11}$$

where $\mathbf{f}(x,t)$ is a $1 \times L$-vector of the (known) f-functions in (x,t) and $\boldsymbol{\beta}$ is a $L \times 1$-vector of the (unknown) β-parameters. The residual term r represents the fluctuations around the trend. The trend vector $\boldsymbol{\beta}$ has a prior distribution given as $N_L(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. In order to do an inversion the posterior distribution for $\boldsymbol{\beta}$ is calculated.

The seismic data $d(x,t)$ in the n locations are represented by the n-vector $\mathbf{d}$. Let $\mathbf{r}$ be the vector of the n residual terms $r(x,t)$ and $\mathbf{e}$ the vector of the n noise terms $e(x,t)$, with

$$\mathbf{e} \sim N_n(\mathbf{0}, \sigma_e^2 \boldsymbol{\Gamma}_e) \quad \text{and} \quad \mathbf{r} \sim N_n(\mathbf{0}, \sigma_r^2 \boldsymbol{\Gamma}_r),$$

where σ_e^2 and σ_r^2 are the variances of $e(x,t)$ and $r(x,t)$ respectively for all (x,t) and $\mathbf{\Gamma}_e$ and $\mathbf{\Gamma}_r$ are the correlation matrices for $\mathbf{e}$ and $\mathbf{r}$ respectively. The Fourier transformed convolutional model from (2) can be written as

$$\begin{aligned} \tilde{d}(\omega,k) &= g(k,\omega)\left[\tilde{\mathbf{f}}(k,\omega)\boldsymbol{\beta}+\tilde{r}(k,\omega)\right]+\tilde{e}(k,\omega) \\ &= g(k,\omega)\tilde{\mathbf{f}}(k,\omega)\boldsymbol{\beta}+\tilde{\epsilon}(k,\omega), \end{aligned} \tag{12}$$

where

$$\tilde{\epsilon}(k,\omega) = g(k,\omega)\tilde{r}(k,\omega)+\tilde{e}(k,\omega). \tag{13}$$

Let $\mathbf{G}$ be a $n \times n$ diagonal matrix with the $g(k,\omega)$ elements on the diagonal, $\tilde{\mathbf{F}}$ a $n \times L$ matrix with the $\tilde{\mathbf{f}}$ vectors on the rows, $\tilde{\boldsymbol{\epsilon}}$ the vector of the $\tilde{\epsilon}(k,\omega)$-terms and $\mathbf{H} = \mathbf{G}\tilde{\mathbf{F}}$. Then expression (12) can be written on vector form as

$$\tilde{\mathbf{d}} = \mathbf{H}\boldsymbol{\beta}+\tilde{\boldsymbol{\epsilon}},$$

where

$$\tilde{\boldsymbol{\epsilon}} = \mathbf{G}\tilde{\mathbf{r}}+\tilde{\mathbf{e}} \sim N_n(\mathbf{0},\tilde{\mathbf{\Sigma}}_\epsilon) \quad \text{and} \quad \tilde{\mathbf{\Sigma}}_\epsilon = \mathbf{G}\tilde{\mathbf{\Sigma}}_r\mathbf{G}^* + \tilde{\mathbf{\Sigma}}_e.$$

Using the theory on circulant covariance matrices described in appendix B.5 gives that

$$\tilde{\mathbf{\Sigma}}_e = \sigma_e^2\tilde{\mathbf{\Lambda}}_e \quad \text{and} \quad \tilde{\mathbf{\Sigma}}_r = \sigma_r^2\tilde{\mathbf{\Lambda}}_r$$

where $\tilde{\mathbf{\Lambda}}_e = \text{diag}\{\tilde{\lambda}_e(k,\omega),\ \forall(k,\omega)\}$ and $\mathbf{\Lambda}_r = \text{diag}\{\tilde{\lambda}_r(k,\omega),\ \forall(k,\omega)\}$ are $n \times n$ diagonal matrices with the eigenvalues for $\mathbf{\Gamma_e}$ and $\mathbf{\Gamma_r}$ respectively on the diagonals, mulitiplied by n. This means that

$$\tilde{\mathbf{\Sigma}}_\epsilon = \text{diag}\{\sigma_r^2\tilde{\lambda}_r(k,\omega)g(k,\omega)g(k,\omega)^* + \sigma_e^2\tilde{\lambda}_e(k,\omega)\}.$$

Straighforward calculations of the posterior (marked by $\star$) mean and covariance for $\boldsymbol{\beta}$ gives

$$\begin{aligned} \boldsymbol{\mu}^\star &= \boldsymbol{\mu}+(\mathbf{\Sigma}\mathbf{H}^*(\mathbf{H}\mathbf{\Sigma}\mathbf{H}^*+\tilde{\mathbf{\Sigma}}_\epsilon)^{-1}(\tilde{\mathbf{d}}-\mathbf{H}\boldsymbol{\mu}) \\ \mathbf{\Sigma}^\star &= \mathbf{\Sigma}-\mathbf{\Sigma}\mathbf{H}^*(\mathbf{H}\mathbf{\Sigma}\mathbf{H}^*+\tilde{\mathbf{\Sigma}}_\epsilon)^{-1}\mathbf{H}\mathbf{\Sigma}. \end{aligned}$$

The matrix to be inverted in these expressions is a $n \times n$-matrix which is typically big. To save computational efforts the posterior mean and covariance can be written as

$$\boldsymbol{\mu}^\star = \boldsymbol{\mu}+(\mathbf{H}^*\tilde{\mathbf{\Sigma}}_\epsilon^{-1}\mathbf{H}+\mathbf{\Sigma}^{-1})^{-1}\mathbf{H}^*\tilde{\mathbf{\Sigma}}_\epsilon^{-1}(\tilde{\mathbf{d}}-\mathbf{H}\boldsymbol{\mu}) \tag{14}$$

$$\mathbf{\Sigma}^\star = (\mathbf{I}-(\mathbf{H}^*\tilde{\mathbf{\Sigma}}_\epsilon^{-1}\mathbf{H}+\mathbf{\Sigma}^{-1})^{-1}\mathbf{H}^*\tilde{\mathbf{\Sigma}}_\epsilon^{-1}\mathbf{H})\mathbf{\Sigma}, \tag{15}$$

which implies inversion of a much smaller $L \times L$-matrix. Expressions (14) and (15) gives an estimate and corresponding uncertainty for $\boldsymbol{\beta}$ which can be used in expression (11) to perform the inversion. The proof for the formulas (14) and (15) is given in appendix C

5.1 Example

The following simple example is constructed on basis of the division of the target area in section 4.1 with the diagonal reflector. This means that $L = 2$ and $f_1(x,t) = 1$ if $(x,t) \in M_1$ and 0 elsewhere and f_2 the opposite. The distribution for $\boldsymbol{\beta}$ is given by

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 1.6 \\ 1.5 \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} 0.0025 & 0 \\ 0 & 0.0025 \end{bmatrix}$$

The covariance structure for the residuals $\mathbf{r}$ is given by the correlation function ν_m in (10) and $\sigma_r^2 = 0.000625$ and the noise term is white noise with variance $\sigma_e^2 = 0.0001$. The synthetic seismic data generated by the convolutional model are shown in Figure 10. The sharp signals in the top and bottom are border effects due to lack of padding.

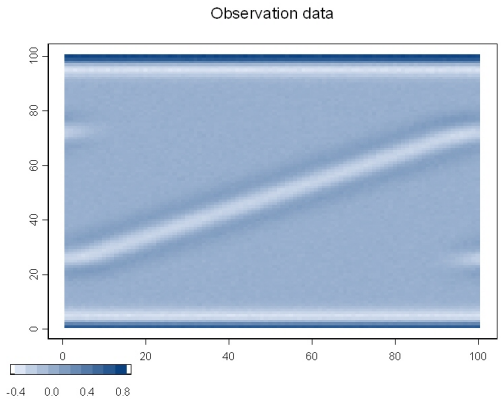


Figure 10. Synthetic seismic data generated by convolutional model with trend

Using formulas (14) and (15) to compute the posterior expectation and covariance for $\boldsymbol{\beta}$ gives

$$\boldsymbol{\mu}^\star = \begin{bmatrix} \mu_1^\star \\ \mu_2^\star \end{bmatrix} = \begin{bmatrix} 1.603217 \\ 1.496783 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Sigma}^\star = \begin{bmatrix} 0.00125222 & 0.00124778 \\ 0.00124778 & 0.00125222 \end{bmatrix} \tag{16}$$

The values for $\boldsymbol{\mu}$ are inserted into the expression $Z_p^\star(x,t) = \exp\left[f_1(x,t)\mu_1 + f_2(x,t)\mu_2\right]$ giving an inverted value for the acoustic impedance. Figure 11 shows these values for the whole target area to the left, and along one trace to the right.

Expression (14) can alternatively be written

$$\boldsymbol{\mu}^\star = [\mathbf{I} - (\mathbf{H}^*\tilde{\boldsymbol{\Sigma}}_\epsilon^{-1}\mathbf{H} + \boldsymbol{\Sigma}^{-1})^{-1}\mathbf{H}^*\tilde{\boldsymbol{\Sigma}}_\epsilon^{-1}\mathbf{H})]\boldsymbol{\mu} + [(\mathbf{H}^*\tilde{\boldsymbol{\Sigma}}_\epsilon^{-1}\mathbf{H} + \boldsymbol{\Sigma}^{-1})^{-1}\mathbf{H}^*\tilde{\boldsymbol{\Sigma}}_\epsilon^{-1}]\tilde{\mathbf{d}}.$$

For the presented model this can be reformulated to

$$\begin{bmatrix} \mu_1^\star \\ \mu_2^\star \end{bmatrix} = \begin{bmatrix} 1-a & a \\ a & 1-a \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} b \\ -b \end{bmatrix} \tag{17}$$

with appropriate values for a and b. From expression (17) it can be deduced that

$$\boldsymbol{\mu}_1^\star + \boldsymbol{\mu}_2^\star = \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2$$

which means that wrong levels in the prior leads to wrong levels in the posterior as seen in the left figure in Figure 11.

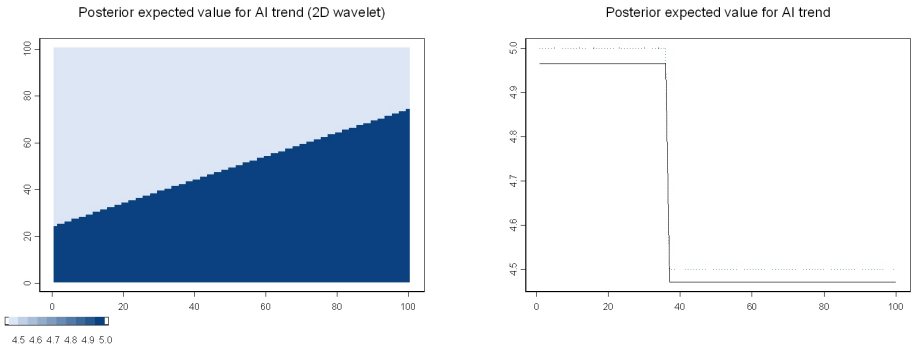


Figure 11. Results of the inversion. Left: Predicted value for the trend in accoustic impedance in the whole target area. Right: Predicted value along one trace.

6 Conclusions

Formulas for inverting seismic data in a 2D target area with a 2D wavelet has been developed and used on simple test cases. The test cases show that for dipping reflectors it seems to be a source to improved inversion results to use a spatial wavelet instead of the traditional 1D deconvolution approach.

References

Aki, K. and Richards, P. G. (1980). *Quantitative seismology: Theory and methods*. W. H. Freeman and Company.

Buland, A., Kolbjørnsen, O., and Omre, H. (2003). Rapid spatially coupled AVO inversion in the Fourier domain. *Geophysics*, 68:824–836.

Dahle, P., Hauge, R., Kolbjørnsen, O., Rossa, E. D., Luoni, F., and Marini, A. (2007). Geostatistical AVO Inversion on a Deep-water Oil Field. Extended Abstract, Petroleum Geostatistics, Cascais, Portugal.

Johnson, R. A. and Wichern, D. W. (1988). *Applied Multivariate Statistical Analysis*. Prentice hall International.

Lecomte, I. (2008). Resolution and illumination analyses in PSDM: A ray-based approach. *The Leading Edge*, May:650–663.

A Fourier transformation theory

A.1 Fourier transform of the derivative

In the expression for the reflection function the derivative of $\tilde{m}(\omega)$ appears. Here it is shown that taking the Fourier transform of the derivative of a function involves multiplying the function itself by $i\omega$.

$$\begin{aligned}
\tilde{m}'(\omega) &= \int \frac{m(t+\Delta t)-m(t)}{\Delta t}\exp\{-i\omega t\}dt \\
&= \frac{1}{\Delta t}\left[\int m(t+\Delta t)\exp\{-i\omega t\}dt - \int m(t)\exp\{-i\omega t\}dt\right] \\
&= \frac{1}{\Delta t}\left[\int m(t+\Delta t)\exp\{-i\omega(t+\Delta t-\Delta t)\}dt - \tilde{m}(\omega)\right] \\
&= \frac{1}{\Delta t}\left[\int m(t+\Delta t)\exp\{-i\omega(t+\Delta t)\}\exp\{i\omega\Delta t\}dt - \tilde{m}(\omega)\right] \\
&= \frac{1}{\Delta t}\left[\tilde{m}(\omega)(\exp\{-i\omega\Delta t\}-1)\right] \\
&= \tilde{m}(\omega)\left[\frac{1}{\Delta t}(\cos(\omega\Delta t)+i\sin(\omega\Delta t)-1)\right] \\
&= \tilde{m}(\omega)\left[\frac{\cos(\omega\Delta t)-1}{\Delta t}+i\omega\frac{\sin(\omega\Delta t)}{\omega\Delta t}\right].
\end{aligned}$$

Taking the limit when $\Delta t \to 0$ gives that the first part inside the paranthesis goes to 0 while the second part converges to $i\omega$ giving

$$\tilde{m}'(\mathbf{k},\omega) = i\omega\tilde{m}(\mathbf{k},\omega).$$

This relation is used in expressions (10) and (11) in Buland et al.

A.2 Discrete Fourier Transformation (DFT)

The one-dimensional DFT of order n, of a sequence $f(j)$ for $j = 0,\ldots,n-1$, can be written

$$\tilde{f}(k) = \sum_{j=0}^{n-1} f(j)\psi^{jk}, \qquad k = 0,\ldots,n-1,$$

where $\psi = \exp\{-2\pi i/n\}$. The inverse transform (IDFT) is then

$$f(j) = \frac{1}{n}\sum_{k=0}^{n-1}\tilde{f}(k)\psi^{-jk}, \qquad j = 0,\ldots,n-1.$$

Alternatively the DFT can be written on matrix form as

$$\tilde{\mathbf{f}} = \mathbf{F}\mathbf{f} \tag{A.1}$$

where $\mathbf{f} = [f(0),\ldots,f(n-1)]^T$ and

$$\mathbf{F} = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ 1 & \psi^1 & \ldots & \psi^{(n-1)} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \psi^{(n-1)} & \ldots & \psi^{(n-1)^2} \end{bmatrix}, \tag{A.2}$$

and the IDFT

$$\mathbf{f} = \frac{1}{n}\mathbf{F}^*\tilde{\mathbf{f}}. \tag{A.3}$$

A.3 Circularity

An important property of the DFT is that it is invariant to a shift by n elements.

$$\begin{aligned}
\tilde{f}(l+n) &= \sum_{k=0}^{n-1} f(k)\exp\{-2\pi i\frac{k(l+n)}{n}\} \\
&= \sum_{k=0}^{n-1} f(k)\exp\{-2\pi i\frac{kl}{n}\}\exp\{-2\pi ik\} \\
&= \sum_{k=0}^{n-1} f(k)\exp\{-2\pi i\frac{kl}{n}\} \\
&= \tilde{f}(l)
\end{aligned}$$

because $\exp\{-2\pi ik\} = i\sin(-2\pi k) + \cos(-2\pi k) = 1$ when k is an integer. This justifies centering the frequency domain around 0. (In $< -n/2, n/2]$)

B Matrix theory

B.1 Cholesky decomposition

If A is a hermitian (symmetric) and positive definite matrix, then A can be written as

$$A = LL^*$$

where L is lower triangular with positive elements on the diagonal.

B.2 Singular value decomposition

If A is a $p \times q$-matrix with rank r, then A can be written as

$$A = ULV^*,$$

where U and V are $p \times r$ and $r \times q$ orthonormal matrices (i.e. $U^*U = V^*V = I$) and L is a $r \times r$-matrix with positive elements on the diagonal and zero elsewhere.

B.3 Spectral decomposition

Every symmetric $p \times p$ matrix A can be written as

$$A = \Gamma^*\Lambda\Gamma = \sum \lambda_i\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i^*, \tag{B.4}$$

where Λ is a diagonal matrix of the eigenvalues $\lambda_1, \ldots, \lambda_p$ of A. The eigenvectors $\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_p$ can be chosen such that $\boldsymbol{\gamma}_i^*\boldsymbol{\gamma}_i = 1$ for all i and $\boldsymbol{\gamma}_i^*\boldsymbol{\gamma}_i = 0$ for all $i \neq j$. Then $\Gamma = [\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_p]$ is an orthonormal matrix with $\Gamma^* = \Gamma^{-1}$ and expression (B.4) is equivalent with

$$\Gamma A\Gamma^* = \Lambda.$$

B.4 Circulant matrices

A $n \times n$ matrix is called circulant when it is on the form

$$\mathbf{C} = \begin{bmatrix} c_0 & c_1 & \dots & c_n - 1 \\ c_n - 1 & c_0 & \dots & c_n - 2 \\ \vdots & \vdots & \vdots & \vdots \\ c_1 & c_2 & \dots & c_0 \end{bmatrix}.$$

The matrix has only n different elements and each row is a cyclic shift of the above row. It can be shown that the eigenvalues of the circulant matrix $\mathbf{C}$ are

$$\lambda_k = \sum_{j=0}^{n-1} c_j \exp\{-2\pi i \frac{jk}{n}\} = \sum_{j=0}^{n-1} c_j \psi^{jk},$$

where $\psi = \exp\{-2\pi i/n\}$ and that the corresponding orthonormal eigenvectors are

$$\mathbf{e}_k = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 \\ \psi^k \\ \dots \\ \psi^{(n-1)k} \end{bmatrix}.$$

Define the eigenvalue matrix $\mathbf{\Lambda}_C = \text{diag}\{\lambda_0, \dots, \lambda_{(n-1)}\}$ and the eigenvector matrix $\mathbf{E}$ as

$$\mathbf{E} = [\mathbf{e}_0, \dots, \mathbf{e}_{n-1}].$$

Then, by performing a spectral decomposition (see section B.3) $\mathbf{C}$ can be diagonalized as

$$\mathbf{E}\mathbf{C}\mathbf{E}^* = \mathbf{\Lambda}_C$$

This means that the eigenvalues of a circulant matrix $\mathbf{C}$ are equal to the DFT of the first row and that the Fourier matrix from A.2 is given by $\mathbf{F} = \sqrt{n}\mathbf{E}$.

B.5 Circulant covariance matrix

Let $\mathbf{Z}(x)$ be a zero-mean Gaussian random field defined on a regular discrete grid, $x_k = k\Delta x$, where $k = 0, \dots, n_x$. The homogenous covariance matrix for $\mathbf{Z}$ is given by

$$\mathbf{\Sigma} = \sigma^2 \begin{bmatrix} \nu_0 & \nu_1 & \cdots & \nu_{n_x-1} \\ \nu_1 & \nu_0 & \cdots & \nu_{n_x-2} \\ \vdots & \vdots & & \vdots \\ \nu_{n_x-1} & \nu_{n_x-2} & \cdots & \nu_0 \end{bmatrix},$$

where $\nu_k = \nu(k\Delta x)$ is the correlation between two gridpoints seperated by k gridcells. Set $n = 2(n_x - 1)$. Then a circulant $n \times n$ matrix $\mathbf{\Sigma}_C$ can be constructed from $\mathbf{\Sigma}$ by the extension

$$\mathbf{\Sigma}_C = \sigma^2 \begin{bmatrix} \nu_0 & \nu_1 & \cdots & \nu_{n_x-1} & \nu_{n_x-2} & \cdots & \nu_1 \\ \nu_1 & \nu_0 & \cdots & \nu_{n_x-2} & \nu_{n_x-1} & \cdots & \nu_2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \nu_{n_x-1} & \nu_{n_x-2} & \cdots & \nu_0 & \nu_1 & \cdots & \nu_{n_x-2} \\ \nu_{n_x-2} & \nu_{n_x-1} & \cdots & \nu_1 & \nu_0 & \cdots & \nu_{n_x-3} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \nu_1 & \nu_2 & \cdots & \nu_{n_x-2} & \nu_{n_x-1} & \cdots & \nu_0 \end{bmatrix},$$

where $\mathbf{\Sigma}$ is the upper left corner of $\mathbf{\Sigma}_C$. This matrix is is a legal covariance matrix if and only if it is positive definite. This can be ensured by truncating the correlation function ν in k_0 where $k_0 < n_x$.

Let $\mathbf{Z}_C$ the the random field with covariance matrix $\mathbf{\Sigma}_C$. The Fourier transform of $\mathbf{Z}_C$, $\tilde{\mathbf{Z}} = \mathbf{F}\mathbf{Z}_C$, has covariance matrix

$$\tilde{\mathbf{\Sigma}}_C = \mathbf{F}\mathbf{\Sigma}_C\mathbf{F}^* = \sqrt{n}\mathbf{E}_C\mathbf{\Sigma}_C\sqrt{n}\mathbf{E}_C^* = n\mathbf{E}_C\mathbf{\Sigma}_C\mathbf{E}_C^{-1} = n\mathbf{\Lambda}_C, \tag{B.5}$$

where $\mathbf{E}_C$ is the matrix of the orthonormal eigenvectors of $\mathbf{\Sigma}_C$ (so $\mathbf{E}_C^* = \mathbf{E}_C^{-1}$) and $\mathbf{\Lambda}_C$ is the diagonal matrix of the (real, nonnegative) eigenvalues of $\mathbf{\Sigma}_C$. From above we know that $\mathbf{\Lambda}_C$ is calculated from the DFT of the first row in $\mathbf{\Sigma}_C$. This means that this operation is enough to specify the distribution for $\mathbf{Z}_C$ of n independent Gaussian variables, and thereby also for the n_x independent variables in $\mathbf{Z}$.

C Proof of expression for posterior mean and covariance for trend coefficient

Setting $\tilde{\boldsymbol{\epsilon}'} = \tilde{\mathbf{\Sigma}}_\epsilon^{-1/2}\tilde{\boldsymbol{\epsilon}}$ gives that $\boldsymbol{\epsilon}' \sim N_n(\mathbf{0}, \mathbf{I})$. Let $\mathbf{L}$ be the lower triangular Cholesky decomposition matrix of $\mathbf{\Sigma}$ (see section B.1), i.e. $\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^*$, set $\boldsymbol{\beta}' = \mathbf{L}^*\mathbf{\Sigma}^{-1}\boldsymbol{\beta}$ and consider the transformation

$$\tilde{\mathbf{d}}' = \tilde{\mathbf{\Sigma}}_\epsilon^{-1/2}\tilde{\mathbf{d}} = \tilde{\mathbf{\Sigma}}_\epsilon^{-1/2}(\mathbf{H}\boldsymbol{\beta} + \tilde{\boldsymbol{\epsilon}}) = \tilde{\mathbf{\Sigma}}_\epsilon^{-1/2}\mathbf{H}\mathbf{L}\boldsymbol{\beta}' + \tilde{\boldsymbol{\epsilon}'} = \mathbf{K}\boldsymbol{\beta}' + \tilde{\boldsymbol{\epsilon}'}$$

where $\mathbf{K} = \tilde{\mathbf{\Sigma}}_\epsilon^{-1/2}\mathbf{H}\mathbf{L}$. This gives

$$\begin{aligned}
\mathrm{Cov}(\boldsymbol{\beta}') &= \mathrm{Cov}(\mathbf{L}^*\mathbf{\Sigma}^{-1}\boldsymbol{\beta}) = \mathbf{L}^*\mathbf{\Sigma}^{-1}\mathbf{\Sigma}\mathbf{\Sigma}^{-*}\mathbf{L} = \mathbf{I} \\
\mathrm{Cov}(\boldsymbol{\beta}', \tilde{\mathbf{d}}') &= \mathrm{Cov}(\boldsymbol{\beta}', \mathbf{K}\boldsymbol{\beta}') = \mathrm{Cov}(\boldsymbol{\beta}')\mathbf{K}^* = \mathbf{K}^* \\
\mathrm{Cov}(\tilde{\mathbf{d}}') &= \mathrm{Cov}(\mathbf{K}\boldsymbol{\beta}') + \mathrm{Cov}(\tilde{\boldsymbol{\epsilon}'}) = \mathbf{K}\mathbf{K}^* + \mathbf{I},
\end{aligned}$$

and by using formulas for multinormal conditional distributions, the posterior mean and covariance for $\boldsymbol{\beta}'$ are

$$\mathrm{E}^\star(\boldsymbol{\beta}') = \mathrm{E}(\boldsymbol{\beta}') + \mathbf{K}^*(\mathbf{K}\mathbf{K}^* + \mathbf{I})^{-1}(\tilde{\mathbf{d}}' - \mathrm{E}(\tilde{\mathbf{d}}')) \tag{C.6}$$

$$\mathrm{Cov}^\star(\boldsymbol{\beta}') = \mathbf{I} - \mathbf{K}^*(\mathbf{K}\mathbf{K}^* + \mathbf{I})^{-1}\mathbf{K} \tag{C.7}$$

In the expressions (C.6) and (C.7) the term $\mathbf{K}^*(\mathbf{K}\mathbf{K}^* + \mathbf{I})^{-1}$ which involves inversion of a $n \times n$-matrix appears. Since n is typically large, and much larger than L, consider the relation

$$\mathbf{K}^*(\mathbf{K}\mathbf{K}^* + \mathbf{I})^{-1} = (\mathbf{K}^*\mathbf{K} + \mathbf{I})^{-1}\mathbf{K}^*. \tag{C.8}$$

which involves inversion of a $L \times L$-matrix. To prove the relation (C.8), consider the singular value decomposition (see section B.2)

$$\mathbf{K} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^*$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices and $\mathbf{\Lambda}$ is a diagonal matrix with positive elements on the diagonal. This gives

$$(\mathbf{K}\mathbf{K}^* + \mathbf{I})^{-1} = (\mathbf{U}(\mathbf{\Lambda}^2\mathbf{U}^* + \mathbf{U}^*))^{-1} = ((\mathbf{\Lambda}^2 + \mathbf{I})\mathbf{U}^*)^{-1}\mathbf{U}^* = \mathbf{U}(\mathbf{\Lambda}^2 + \mathbf{I})^{-1}\mathbf{U}^*,$$

so

$$\mathbf{K}^*(\mathbf{K}\mathbf{K}^* + \mathbf{I})^{-1} = \mathbf{V}\mathbf{\Lambda}(\mathbf{\Lambda}^2 + \mathbf{I})^{-1}\mathbf{U}^*. \tag{C.9}$$

On the other hand,

$$(\mathbf{K}^*\mathbf{K} + \mathbf{I})^{-1} = (\mathbf{V}(\mathbf{\Lambda}^2\mathbf{V}^* + \mathbf{V}^*))^{-1} = \mathbf{V}(\mathbf{\Lambda}^2 + \mathbf{I})^{-1}\mathbf{V}^*,$$

so

$$(\mathbf{K}^*\mathbf{K} + \mathbf{I})^{-1}\mathbf{K}^* = \mathbf{V}(\mathbf{\Lambda}^2 + \mathbf{I})^{-1}\mathbf{\Lambda}\mathbf{U}^*$$

which equals expression (C.9) since $\mathbf{\Lambda}$ and $(\mathbf{\Lambda}^2 + \mathbf{I})^{-1}$ are both diagonal matrices and thereby $(\mathbf{\Lambda}^2 + \mathbf{I})^{-1}\mathbf{\Lambda} = \mathbf{\Lambda}(\mathbf{\Lambda}^2 + \mathbf{I})^{-1}$. Using $\mathbf{K} = \tilde{\mathbf{\Sigma}}_\epsilon^{-1/2}\mathbf{H}\mathbf{L}$ gives

$$\begin{aligned}
\mathbf{K}^*(\mathbf{K}\mathbf{K}^* + \mathbf{I})^{-1} &= (\mathbf{K}^*\mathbf{K} + \mathbf{I})^{-1}\mathbf{K}^* \\
&= (\mathbf{L}^*(\mathbf{H}^*\tilde{\mathbf{\Sigma}}_\epsilon^{-1}\mathbf{H}\mathbf{L} + \mathbf{\Sigma}^{-1}\mathbf{L}))^{-1}\mathbf{L}^*\mathbf{H}^*\tilde{\mathbf{\Sigma}}_\epsilon^{-1/2} \\
&= ((\mathbf{H}^*\tilde{\mathbf{\Sigma}}_\epsilon^{-1}\mathbf{H} + \mathbf{\Sigma}^{-1})\mathbf{L})^{-1}\mathbf{H}^*\tilde{\mathbf{\Sigma}}_\epsilon^{-1/2} \\
&= \mathbf{L}^{-1}(\mathbf{H}^*\tilde{\mathbf{\Sigma}}_\epsilon^{-1}\mathbf{H} + \mathbf{\Sigma}^{-1})^{-1}\mathbf{H}^*\tilde{\mathbf{\Sigma}}_\epsilon^{-1/2}.
\end{aligned}$$

Applying this to expressions (C.6) and (C.7) gives

$$\begin{aligned}
\mathrm{E}^\star(\boldsymbol{\beta}') &= \mathbf{L}^*\mathbf{\Sigma}^{-1}\boldsymbol{\mu} + \mathbf{L}^{-1}(\mathbf{H}^*\tilde{\mathbf{\Sigma}}_\epsilon^{-1}\mathbf{H} + \mathbf{\Sigma}^{-1})^{-1}\mathbf{H}^*\tilde{\mathbf{\Sigma}}_\epsilon^{-1/2}(\tilde{\mathbf{d}}' - \mathrm{E}(\tilde{\mathbf{d}}')) \\
\mathrm{Cov}^\star(\boldsymbol{\beta}') &= \mathbf{I} - \mathbf{L}^{-1}(\mathbf{H}^*\tilde{\mathbf{\Sigma}}_\epsilon^{-1}\mathbf{H} + \mathbf{\Sigma}^{-1})^{-1}\mathbf{H}^*\tilde{\mathbf{\Sigma}}_\epsilon^{-1}\mathbf{H}\mathbf{L}
\end{aligned}$$

Since $\boldsymbol{\beta} = \mathbf{\Sigma}\mathbf{L}^{-*}\boldsymbol{\beta}'$,

$$\begin{aligned}
\boldsymbol{\mu}^\star = \mathrm{E}^\star(\boldsymbol{\beta}) &= \boldsymbol{\mu} + (\mathbf{H}^*\mathbf{D}^{-1}\mathbf{H} + \mathbf{\Sigma}^{-1})^{-1}\mathbf{H}^*\mathbf{D}^{-1}(\tilde{\mathbf{d}} - \mathbf{H}\boldsymbol{\mu}) \\
\mathbf{\Sigma}^\star = \mathrm{Cov}^\star(\boldsymbol{\beta}) &= \mathbf{\Sigma}\mathbf{L}^{-*}\left[\mathbf{I} - \mathbf{L}^{-1}(\mathbf{H}^*\mathbf{D}^{-1}\mathbf{H} + \mathbf{\Sigma}^{-1})^{-1}\mathbf{H}^*\mathbf{D}^{-1}\mathbf{H}\mathbf{L}\right]\mathbf{L}^{-1}\mathbf{\Sigma} \\
&= (\mathbf{I} - (\mathbf{H}^*\mathbf{D}^{-1}\mathbf{H} + \mathbf{\Sigma}^{-1})^{-1}\mathbf{H}^*\mathbf{D}^{-1}\mathbf{H})\mathbf{\Sigma}.
\end{aligned}$$