Estimation of Smoothness from the Residual Field

FMRIB Technical Report TR00MJ3

Mark Jenkinson

Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB),
Department of Clinical Neurology, University of Oxford, John Radcliffe Hospital,
Headley Way, Headington, Oxford, UK

Abstract

This report gives a detailed derivation of the smoothness estimation techniques used in Kiebel99 [2] and Forman95 [1].

1 Introduction

The task is to find a statistic that will estimate the smoothing parameters of a Gaussian Random Field. Such a field can be constructed by taking an uncorrelated, zero-mean (white-noise) field, F_W, and smoothing it spatially with a Gaussian filter. That is, the smoothed field F_S is given by:

$$F_S(\mathbf{x}) = G(\mathbf{x}) \otimes F_W(\mathbf{x}) = \int G(\mathbf{x}-\mathbf{p}) F_W(\mathbf{p})\, d\mathbf{p} \tag{1}$$

where $\otimes$ denotes convolution and

$$\begin{aligned} G(\mathbf{x}) &= G_{\sigma_x}(x)\, G_{\sigma_y}(y)\, G_{\sigma_z}(z) && (2)\\ G_\sigma(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp(\frac{-x^2}{2\sigma^2}) && (3)\end{aligned}$$

is the spatial filter. Note that all integrals in this report are definite integrals over all space (that is, from $-\infty$ to $+\infty$ in 1D).

In addition, the auto-correlation of F_W is:

$$E\{F_W(\mathbf{x_1})F_W(\mathbf{x_2})\} = \delta(\mathbf{x_1} - \mathbf{x_2}) \tag{4}$$

where $\delta(\cdots)$ is the Dirac delta function.

2 Covariance of the Smoothed Field

Consider the covariance of the smoothed field:

$$\begin{aligned} E\{F_S(\mathbf{x_1})F_S(\mathbf{x_2})\} &= E\left\{\int\int G(\mathbf{x_1}-\mathbf{p_1})G(\mathbf{x_2}-\mathbf{p_2})F_W(\mathbf{p_1})F_W(\mathbf{p_2})\, d\mathbf{p_1}\, d\mathbf{p_2}\right\} && (5)\\ &= \int\int G(\mathbf{x_1}-\mathbf{p_1})G(\mathbf{x_2}-\mathbf{p_2})\delta(\mathbf{p_1}-\mathbf{p_2})\, d\mathbf{p_1}\, d\mathbf{p_2} && (6)\\ &= \int G(\mathbf{x_1}-\mathbf{x_0})G(\mathbf{x_2}-\mathbf{x_0})\, d\mathbf{x_0} && (7)\\ &= \left(\int G_{\sigma_x}(x_1-x_0)G_{\sigma_x}(x_2-x_0)\, dx_0\right)\left(\int G_{\sigma_y}(y_1-y_0)G_{\sigma_y}(y_2-y_0)\, dy_0\right) \times \cdots \\ &\qquad \left(\int G_{\sigma_z}(z_1-z_0)G_{\sigma_z}(z_2-z_0)\, dz_0\right) && (8)\end{aligned}$$

Each 1D integral is of the form:

$$
\begin{aligned}
I_1 &= \int G_\sigma(x_1 - x_0) G_\sigma(x_2 - x_0)\, dx_0 && (9)\\
&= \frac{1}{2\pi\sigma^2} \int \exp\left(\frac{-1}{2\sigma^2} \left[(x_1 - x_0)^2 + (x_2 - x_0)^2 \right] \right) dx_0 && (10)\\
&= \frac{1}{2\pi\sigma^2} \int \exp\left(\frac{-1}{2\sigma^2} \frac{(x_1 - x_2)^2}{2} \right) \exp\left(\frac{-1}{\sigma^2} \left[x_0 - \frac{x_1 + x_2}{2} \right]^2 \right) dx_0 && (11)\\
&= \frac{1}{2\pi\sigma^2} \exp\left(\frac{-(x_1 - x_2)^2}{4\sigma^2} \right) \sqrt{\pi\sigma^2} && (12)\\
&= \frac{1}{\sqrt{4\pi\sigma^2}} \exp\left(\frac{-(x_1 - x_2)^2}{4\sigma^2} \right) && (13)
\end{aligned}
$$

Therefore:

$$
E\{F_S(\mathbf{x_1})F_S(\mathbf{x_2})\} = \frac{1}{(4\pi)^{\frac{3}{2}}\sigma_x\sigma_y\sigma_z} \exp\left(\frac{-(x_1 - x_2)^2}{4{\sigma_x}^2} \right) \exp\left(\frac{-(y_1 - y_2)^2}{4{\sigma_y}^2} \right) \exp\left(\frac{-(z_1 - z_2)^2}{4{\sigma_z}^2} \right) \quad (14)
$$

3 Covariance of the Smoothed Derivative Field

Consider taking the spatial derivative of the smoothed field. That is, constructing three derivative fields, F_{Dx}, F_{Dy}, F_{Dz}, where:

$$
F_{Dx}(\mathbf{x}) = \frac{\partial}{\partial x} F_S(\mathbf{x}) = \left(\frac{\partial G(\mathbf{x})}{\partial x} \right) \otimes F_W(\mathbf{x}) \quad (15)
$$

and similarly for F_{Dy} and F_{Dz}.

Now the partial derivative of the smoothing filter is given by:

$$
\frac{\partial G(\mathbf{x})}{\partial x} = \frac{dG_{\sigma_x}(x)}{dx} G_{\sigma_y}(y) G_{\sigma_z}(z) = \frac{-x}{{\sigma_x}^2} G_{\sigma_x}(x) G_{\sigma_y}(y) G_{\sigma_z}(z). \quad (16)
$$

Therefore, the covariance can be computed as in the last section:

$$
\begin{aligned}
E\{F_{Dx}(\mathbf{x_1})F_{Dx}(\mathbf{x_2})\} &= E\left\{ \int\int \frac{\partial G(\mathbf{x_1} - \mathbf{p_1})}{\partial x} \frac{\partial G(\mathbf{x_2} - \mathbf{p_2})}{\partial x} F_W(\mathbf{p_1}) F_W(\mathbf{p_2})\, d\mathbf{p_1}\, d\mathbf{p_2} \right\} && (17)\\
&= \int\int \frac{\partial G(\mathbf{x_1} - \mathbf{p_1})}{\partial x} \frac{\partial G(\mathbf{x_2} - \mathbf{p_2})}{\partial x} \delta(\mathbf{p_1} - \mathbf{p_2})\, d\mathbf{p_1}\, d\mathbf{p_2} && (18)\\
&= \int \frac{\partial G(\mathbf{x_1} - \mathbf{x_0})}{\partial x} \frac{\partial G(\mathbf{x_2} - \mathbf{x_0})}{\partial x} d\mathbf{x_0} && (19)\\
&= \left(\int \frac{(x_1 - x_0)(x_2 - x_0)}{{\sigma_x}^4} G_{\sigma_x}(x_1 - x_0) G_{\sigma_x}(x_2 - x_0)\, dx_0 \right) \times \cdots \\
&\quad \left(\int G_{\sigma_y}(y_1 - y_0) G_{\sigma_y}(y_2 - y_0)\, dy_0 \right) \left(\int G_{\sigma_z}(z_1 - z_0) G_{\sigma_z}(z_2 - z_0)\, dz_0 \right) && (20)
\end{aligned}
$$

The latter two integrals are given by equation 13. The former integral is:

$$
\begin{aligned}
I_2 &= \int \frac{(x_1 - x_0)(x_2 - x_0)}{\sigma^4} G_\sigma(x_1 - x_0) G_\sigma(x_2 - x_0)\, dx_0 && (21)\\
&= \frac{1}{2\pi\sigma^6} \exp\left(\frac{-(x_1 - x_2)^2}{4\sigma^2} \right) \int (x_1 - x_0)(x_2 - x_0) \exp\left(\frac{-1}{\sigma^2} \left[x_0 - \frac{x_1 + x_2}{2} \right]^2 \right) dx_0 && (22)
\end{aligned}
$$

Rewriting the last part using $w = x_0 - \frac{x_1 + x_2}{2}$ gives:

$$
I_3 = \int (x_1 - x_0)(x_2 - x_0) \exp\left(\frac{-1}{\sigma^2} \left[x_0 - \frac{x_1 + x_2}{2} \right]^2 \right) dx_0 \quad (23)
$$

$$= \int \left(w - \frac{x_1 - x_2}{2}\right)\left(w + \frac{x_1 - x_2}{2}\right) \exp\left(\frac{-w^2}{\sigma^2}\right) dw \quad (24)$$

$$= \int w^2 \exp\left(\frac{-w^2}{\sigma^2}\right) dw - \frac{(x_1 - x_2)^2}{4} \int \exp\left(\frac{-w^2}{\sigma^2}\right) dw \quad (25)$$

$$= \frac{\sigma^2}{2}\sqrt{\pi\sigma^2} - \frac{(x_1 - x_2)^2}{4}\sqrt{\pi\sigma^2} \quad (26)$$

$$= (\pi)^{\frac{1}{2}}\sigma\left(\frac{\sigma^2}{2} - \frac{(x_1 - x_2)^2}{4}\right) \quad (27)$$

Therefore, combining equations 20, 22 and 27 gives:

$$E\{F_{Dx}(\mathbf{x_1})F_{Dx}(\mathbf{x_2})\} = E\{F_S(\mathbf{x_1})F_S(\mathbf{x_2})\}\frac{1}{\sigma_x{}^4}\left(\frac{\sigma_x{}^2}{2} - \frac{(x_1 - x_2)^2}{4}\right) \quad (28)$$

4 Normalisation

Consider a scalar multiple of a field. If this scalar is constant and independent of position, then the covariance of the field scales with the constant squared. For instance, let $S(\mathbf{x}) = kF(\mathbf{x})$, then

$$E\{S(\mathbf{x_1})S(\mathbf{x_2})\} = k^2 E\{F(\mathbf{x_1})F(\mathbf{x_2})\}. \quad (29)$$

The importance of this result is that by choosing an appropriate scaling factor, the covariance of the new smoothed field, S, can be set to any constant value — in agreement with the experimentally normalised results.

5 Estimation of Kiebel *et al.* (SPM)

This section describes the theory and practice of the robust smoothness estimation described in [2]. The link between the notation used in the previous sections and the notation used in the paper is also made explicit.

5.1 Theoretical Basis

To estimate the smoothing variances, $\sigma_x{}^2, \sigma_y{}^2, \sigma_z{}^2$, the expectation of the partial derivative is used. That is,

$$E\left\{\left(\frac{\partial S}{\partial x}\right)^2\right\} = \frac{1}{2\sigma_x{}^2} \quad (30)$$

where the quantity S is a scaled version of F_S such that $E\{S^2\} = 1$.

This follows easily from equations 28 and 14 by setting $\mathbf{x_1} = \mathbf{x_2} = \mathbf{x}$. That is:

$$E\{(F_{Dx}(\mathbf{x}))^2\} = E\{(F_S(\mathbf{x}))^2\}\left(\frac{1}{2\sigma_x{}^2}\right) \quad (31)$$

$$E\{(F_S(\mathbf{x}))^2\} = \frac{1}{(4\pi)^{\frac{3}{2}}\sigma_x\sigma_y\sigma_z}. \quad (32)$$

Therefore, by setting $S(\mathbf{x}) = kF_S(\mathbf{x})$ with $k = (4\pi)^{\frac{3}{4}}(\sigma_x\sigma_y\sigma_z)^{\frac{1}{2}}$ gives $E\{S^2(\mathbf{x})\} = 1$ and $\frac{\partial S(\mathbf{x})}{\partial x} = kF_{Dx}(\mathbf{x})$, which leads to equation 30.

5.2 Sampling Statistics

In practice, there are a finite number of samples of the residual field that are available. This takes the form of a regular 4D array of samples, including the three spatial dimensions and the temporal dimension. These samples (post-normalisation) shall be denoted as: $S_t(\mathbf{x})$ where t refers to the time index and $\mathbf{x}$ takes discrete values. The number of samples in each dimension is N_X by N_Y by N_Z by N_T, with $N = N_X N_Y N_Z$ being a shorthand for the number of voxels.

Now, each sample point in the 4D field is a random variable, with expectation given by equation 30. Therefore, due to the linearity of the expectation, the results can be averaged over all the sample points to achieve a more accurate estimate. That is, for *a single point in time only*:

$$\lambda_{11} = \frac{1}{N} \sum_{\mathbf{x}} \left(\frac{\partial S(\mathbf{x})}{\partial x} \right)^2 \tag{33}$$

such that

$$E\{\lambda_{11}\} = \frac{1}{N} \sum_{\mathbf{x}} E\left\{ \left(\frac{\partial S(\mathbf{x})}{\partial x} \right)^2 \right\} \tag{34}$$

$$= \frac{1}{N} \sum_{\mathbf{x}} \frac{1}{2{\sigma_x}^2} \tag{35}$$

$$= \frac{1}{2{\sigma_x}^2}. \tag{36}$$

Similarly for λ_{22} and λ_{33}.

Note that $E\{\lambda_{11}\} = \frac{1}{2{\sigma_x}^2}$, $E\{\lambda_{22}\} = \frac{1}{2{\sigma_y}^2}$ and $E\{\lambda_{33}\} = \frac{1}{2{\sigma_z}^2}$ is the same notation used in [2].

5.3 Normalisation

Let an unnormalised voxel time series be denoted by $R_t(\mathbf{x})$ where $t = 1, \ldots, N_T$ is the time index. Treating this as a time-vector (that is, an N_T by 1 matrix), SPM performs a "normalisation":

$$S_t(\mathbf{x}) = \frac{R_t(\mathbf{x})}{\sqrt{\sum_{t=1}^{N_T} R_t(\mathbf{x}) R_t(\mathbf{x})}} \tag{37}$$

or, by suppressing indices and using matrix notation, as $S = R/\sqrt{R^\top R}$.

This "normalisation" results in the expected *sum* of the residuals squared being unity. Therefore, the expected value for any particular residual squared is actually:

$$E\{(S_t(\mathbf{x})\,)^2\} = \frac{1}{N_T}. \tag{38}$$

Consequently, the factor k, introduced previously needs to be divided by $\sqrt{N_T}$. This also means that, when taking the sum over the possible time points, it is no longer necessary to normalise the sum. Therefore, the full 4D average for λ, assuming *no temporal correlation*, is:

$$\lambda_{11} = \frac{1}{N} \sum_{\mathbf{x}} \sum_{t=1}^{N_T} \left(\frac{\partial S_t(\mathbf{x})}{\partial x} \right)^2. \tag{39}$$

5.4 Estimating the Derivative

As no direct measurements of the continuous derivative exist, it must be estimated from the samples. This is done using a simple difference form:

$$S_t'(\mathbf{x}) = \frac{S_t(\mathbf{x} + \mathbf{d}) - S_t(\mathbf{x} - \mathbf{d})}{2\|\mathbf{d}\|}. \tag{40}$$

The direction of the derivative is specified by the direction of the difference vector $\mathbf{d}$.

It is assumed that this approximation ($S_t'(\mathbf{x}) \approx \frac{\partial S_t(\mathbf{x})}{\partial x}$) is sufficiently accurate so that estimation described in equation 36 is not significantly biased.

5.5 Summary

In practice the smoothness is estimated using the following steps:

1. Normalise the 4D residuals, $R_t(\mathbf{x}) \to S_t(\mathbf{x})$, using equation 37.

2. Calculate the derivative approximation $S'_t(\mathbf{x})$ at all 4D samples using equation 40.
3. Calculate the 4D average statistic λ_{jj} using equation 39.
4. Accumulate the results of λ_{jk} into a matrix (assuming non-diagonal elements are zero):

$$\Lambda = \begin{pmatrix} \frac{1}{2{\sigma_x}^2} & 0 & 0 \\ 0 & \frac{1}{2{\sigma_y}^2} & 0 \\ 0 & 0 & \frac{1}{2{\sigma_z}^2} \end{pmatrix} \tag{41}$$

5. Optional: Calculate the filter width matrix $W = (2\Lambda)^{-1}$ and the FWHM values as $\text{FWHM}_j = \sqrt{8\ln(2)W_{jj}}$.

6 Estimation of Forman *et al.*

This section describes the theory and practice of using the smoothness estimation presented in [1]. As much of this is similar to the SPM method outlined in the previous section, only the main differences will be highlighted.

6.1 Theoretical Basis

As the approximation of the derivative (equation 40) is not precise, it introduces a bias into the SPM estimation technique. This source of bias can be eliminated by using a finite difference calculation instead of the derivative-based one.

That is, the difference field is:

$$F_\Delta(\mathbf{x}) = F_S(\mathbf{x}+\mathbf{d}) - F_S(\mathbf{x}). \tag{42}$$

The variance of this difference field is:

$$\begin{aligned} E\{F_\Delta(\mathbf{x})F_\Delta(\mathbf{x})\} &= E\{F_S(\mathbf{x}+\mathbf{d})F_S(\mathbf{x}+\mathbf{d}) + F_S(\mathbf{x})F_S(\mathbf{x}) - 2F_S(\mathbf{x}+\mathbf{d})F_S(\mathbf{x})\} & (43) \\ &= 2E\{F_S(\mathbf{x})F_S(\mathbf{x})\} - 2E\{F_S(\mathbf{x}+\mathbf{d})F_S(\mathbf{x})\} & (44) \end{aligned}$$

Setting $\mathbf{x_2} = \mathbf{x_1} + \mathbf{d}_x$ with $\mathbf{d}_x = (d, 0, 0)$, the expectations can be calculated using equation 14:

$$\begin{aligned} E\{F_S(\mathbf{x})F_S(\mathbf{x}+\mathbf{d}_x)\} &= \frac{1}{(4\pi)^{\frac{3}{2}}\sigma_x\sigma_y\sigma_z}\exp\left(\frac{-d^2}{4{\sigma_x}^2}\right) & (45) \\ E\{F_S(\mathbf{x})F_S(\mathbf{x})\} &= \frac{1}{(4\pi)^{\frac{3}{2}}\sigma_x\sigma_y\sigma_z}. & (46) \end{aligned}$$

giving

$$E\{F_\Delta(\mathbf{x})F_\Delta(\mathbf{x})\} = \frac{2}{(4\pi)^{\frac{3}{2}}\sigma_x\sigma_y\sigma_z}\left(1 - \exp\left(\frac{-d^2}{4{\sigma_x}^2}\right)\right) \tag{47}$$

This enables the individual smoothing parameters, $\sigma_x, \sigma_y, \sigma_z$ to be found. That is:

$${\sigma_x}^2 = \frac{-d^2}{4\ln\left(1 - \frac{E\{F_\Delta(\mathbf{x})F_\Delta(\mathbf{x})\}}{2E\{F_S(\mathbf{x})F_S(\mathbf{x})\}}\right)}. \tag{48}$$

6.1.1 A Simple Alternative

Instead of taking the square of the difference, the first order correlation can be used directly to give:

$${\sigma_x}^2 = \frac{-d^2}{4\ln\left(\frac{E\{F_S(\mathbf{x})F_S(\mathbf{x}+\mathbf{d}_x)\}}{E\{F_S(\mathbf{x})F_S(\mathbf{x})\}}\right)} \tag{49}$$

6.2 Sampling Statistics

To estimate the required variances simple averages are taken over the required quantities. That is:

$$
\begin{aligned}
V_0 &= \frac{1}{N}\sum_{\mathbf{x}}\frac{1}{N_T}\sum_{t=1}^{N_T}\left(S_t(\mathbf{x})\right)^2 && (50)\\
V_1 &= \frac{1}{N}\sum_{\mathbf{x}}\frac{1}{N_T}\sum_{t=1}^{N_T}\left(S_t(\mathbf{x}+\mathbf{d})-S_t(\mathbf{x})\right)^2 && (51)\\
V_2 &= \frac{1}{N}\sum_{\mathbf{x}}\frac{1}{N_T}\sum_{t=1}^{N_T}\left(S_t(\mathbf{x}+\mathbf{d})S_t(\mathbf{x})\right) && (52)
\end{aligned}
$$

where $S_t(\mathbf{x})$ is the observed smooth field at time t.

The required smoothing parameters are then calculated using equations 48 and 49. That is:

$$
{\sigma_x}^2 = \frac{-d^2}{4\ln\left(1-\frac{V_1}{2V_0}\right)}. \tag{53}
$$

or

$$
{\sigma_x}^2 = \frac{-d^2}{4\ln\left(\frac{V_2}{V_0}\right)} \tag{54}
$$

Note that, because ratios are used, any constant scaling factors relating $S(\mathbf{x})$ to $F_S(\mathbf{x})$ will cancel.

6.3 Normalisation

Although an individual voxel normalisation:

$$
S_t(\mathbf{x}) = \frac{\sqrt{N_T}\,R_t(\mathbf{x})}{\sqrt{\sum_{t=1}^{N_T} R_t(\mathbf{x})R_t(\mathbf{x})}} \tag{55}
$$

can be used, the estimation of V_0 above effectively eliminates the need for normalisation.

However, this is only valid if the residual field is stationary, or at least approximately stationary. In this case, the variance is constant across all spatial points. If this is not the case, then the normalisation used by SPM (section 5.3) is more likely to produce a field that is a better approximation to a GRF.

A Gaussian Integrals

$$
\begin{aligned}
\int \exp\left(\frac{-x^2}{2\sigma^2}\right)dx &= \sqrt{2\pi\sigma^2} && (56)\\
\int x\exp\left(\frac{-x^2}{2\sigma^2}\right)dx &= 0 && (57)\\
\int x^2\exp\left(\frac{-x^2}{2\sigma^2}\right)dx &= \sigma^2\sqrt{2\pi\sigma^2} && (58)
\end{aligned}
$$

References

[1] Stephen D. Forman, Jonathon D. Cohen, Mark Fitzgerald, William F. Eddy, Mark A. Mintum, and Douglas C. Noll. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine*, 33:636–647, 1995.

[2] Stefan J. Kiebel, Jean-Baptiste Poline, Karl J. Friston, Andrew P. Holmes, and Keith J. Worsley. Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model. *NeuroImage*, 10:756–766, 1999.