Spatial smoothing of autocorrelations to control the degrees of freedom in fMRI analysis⋆

Keith J. Worsley

Department of Mathematics and Statistics, and Montreal Neurological Institute, McGill University, Montreal, Canada

In the statistical analysis of fMRI data, the parameter of primary interest is the effect of a contrast; of secondary interest is its standard error, and of tertiary interest is the standard error of this standard error, or equivalently, the degrees of freedom (df). In a ReML (Restricted Maximum Likelihood) analysis with AR(p) errors, we show how spatial smoothing of temporal autocorrelations increases the effective df (but not the smoothness of primary or secondary parameter estimates), so that the amount of smoothing can be chosen in advance to achieve a target df, typically 100. This has already been done at the second level of a hierarchical analysis by smoothing the ratio of random to fixed effects variances [1]; we now show how to do it at the first level, by smoothing autocorrelation parameters [2].

Methods

The effective df is derived assuming the true lag j autocorrelations $\rho_j = 0$; we then show (Figure 1) that this is reasonably accurate even if $\rho_j \neq 0$.

Suppose X is the $n \times m$ design matrix of the linear model whose columns are the covariates, and let c be an m-vector of contrasts in those columns whose effects we are interested in. Let

$$x = (x_1, \ldots, x_n)' = X(X'X)^{-1}c \tag{1}$$

be the least-squares contrast in the *observations*, and let τ_j be its lag j autocorrelation

$$\tau_j = \sum_{i=j+1}^{n} x_i x_{i-j} \Big/ \sum_{i=1}^{n} x_i^2. \tag{2}$$

Let $FWHM_{\text{data}}$ be the effective FWHM of the fMRI data, and let $FWHM_{\text{filter}}$ be the FWHM of the Gaussian filter used for spatial smoothing of the temporal autocorrelations. Let

$$f = \left(1 + 2\frac{FWHM^2_{\text{filter}}}{FWHM^2_{\text{data}}}\right)^{-D/2} \tag{3}$$

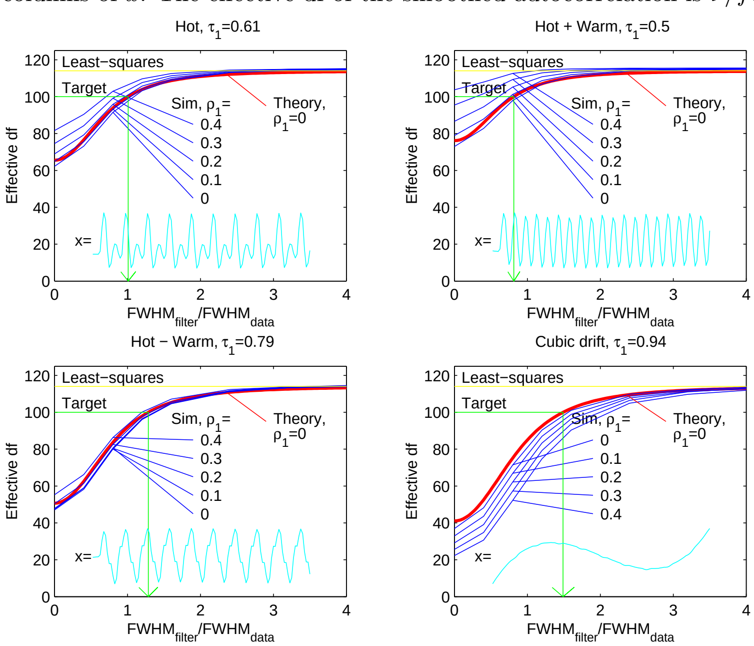
where D is the number of dimensions. Then the effective df of the contrast is

$$\tilde{\nu} \approx \nu / (1 + 2f \sum_{j=1}^{p} \tau_j^2) \tag{4}$$

where $\nu = n - m$ is the usual least-squares residual df. For an F statistic that simultaneously tests k columns of the $m \times k$ contrast matrix C, the effective numerator df is the same as (4) but with x replaced by the normalized matrix of the least-squares contrasts in the observations

$$x = X(X'X)^{-1}C(C'(X'X)^{-1}C)^{-1/2}, \tag{5}$$

so that $x'x$ is the $k \times k$ identity matrix, and with the autocorrelation τ_j replaced by the average of the k temporal autocorrelations of the columns of x. The effective df of the smoothed autocorrelation is ν/f.



Results

Figure 1 (bottom left): Validation of the theoretical effective df $\tilde{\nu}$ (red) with simulations (blue) for AR(1). Four different contrasts were considered; the time course of each contrast in the observations, x, is plotted below in cyan (τ_1 is its lag 1 autocorrelation). Effective df is plotted against the spatial smoothing of the temporal autocorrelations, relative to the spatial smoothing of the fMRI data. Smoothing increases the effective df, up to the least-squares df $\nu = n - m = 120 - 6 = 114$ (yellow). We can see that the theoretical result (4), derived assuming the autocorrelation ρ_1 is zero, is a reasonable approximation even if the autocorrelation is not zero. Using this, we can find the amount of smoothing needed to target a particular df, here 100 (green).

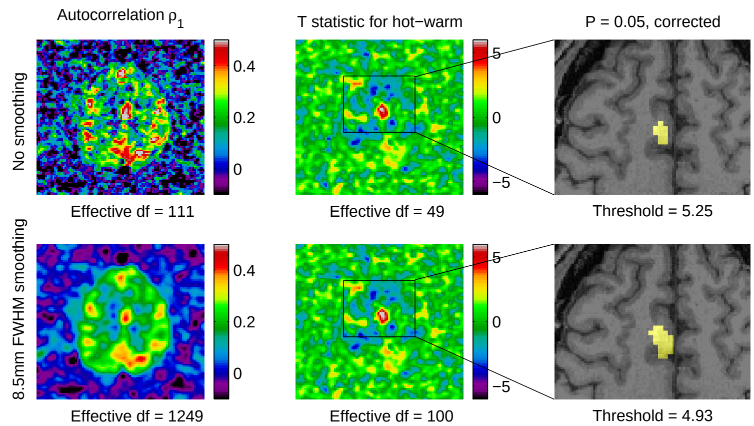


Figure 2: Application to a typical 6 minute, TR=3, 1.5T fMRI data set with 6mm smoothing [1]. Temporal autocorrelation without (top row) and with (bottom row) spatial smoothing, the corresponding T statistics for a hot-warm effect, and the detected activation. 8.5mm smoothing is needed to achieve 100 df, the resulting $P = 0.05$ threshold decreases, and roughly twice as much activation is detected.

Discussion

The proposed method is extremely fast and it does not require any image processing. It can be used in conjunction with other regularization methods [3] to calculate their effective df.

Temporal correlation of the covariates decreases effective df, but since $f \leq 1$, spatial smoothing ameliorates this effect. Reversing this formula (4), we can calculate the amount of smoothing required to achieve a desired df. Note that this will depend on the contrast, so we suggest being conservative by taking the maximum of the amounts of smoothing.

We can never get more than the least-squares df without smoothing the residual variance as well, in which case the factor f would be applied to all the terms in the denominator of (4). Of course we do not wish to do this because the residual variance contains too much anatomical structure, and so smoothing could result in serious biases.

This strategy may not result in any smoothing at all; if in the example in Figure 2 we had twice as many observations over a 12 minute time period then the effective df would be more than 100 without any smoothing, so no smoothing would be necessary. If on the other hand we had a short sequence with less than 100 least-squares df, then no amount of smoothing can increase this above 100 (without smoothing the variance itself). In this case, we recommend choosing the amount of smoothing to target a high proportion, say 90%, of the least-squares df. This strategy has been implemented in the `FMRISTAT` and `BRAINSTAT` packages [4].

References

[1]. Worsley, K.J. *et al.* (2002). *NeuroImage*, **15**:1-15.
[2]. Worsley, K.J. (2005). *Neuroimage*, **26**:635-641.
[3]. Gautama, T. & Van Hulle, M.M. (2004). *NeuroImage*, **23**:1023-1216.
[4]. Taylor, J.E. (2005). *NeuroImage*, Poster #763 T-AM.

⋆#610 M-PM: poster at http://www.math.mcgill.ca/keith/fmristat/poster_ardf.pdf June 13, 2005