Computing Space and Time Autocorrelated Standard Errors

Jordan Adamson, https://github.com/Jadamso/STrollR

Standard errors can be easily corrected for autocorrelation over space and time with sparse matrix techniques. Denote Σ_i as the autocorrelation over time within a cell, and Σ_t as the autocorrelation over space within a time period. Then the covariance matrix for the entire set can be written as separable, $\Sigma = \Sigma_I + \Sigma_T$, where Σ_I denotes Σ_i for all individuals and Σ_T denotes Σ_t for all time periods.[1] To estimate Σ_t, one can use a weight $k_{i,j}$ between observations i and j based on their distance. For example, the Bartlett kernel, also known as the tent or triangle kernel, can create very sparse covariance matrices, as $k_{i,j} = 0$ for many i, j. To exploit this property, I first formulate the 'meat' of the variance estimator,

$$X'K \circ \hat{e}\hat{e}'X = X'K \circ (\mathbf{1} * \hat{e}) \circ (\mathbf{1} * \hat{e})'X = X'(\hat{e} * (K * \hat{e})')'X,$$

where $'$ denotes transpose, K denotes the sparse weights matrix, $\circ$ denotes the Hadamard product of matrices, $*$ denotes the column-wise Hadamard product of a matrix and a vector, and $\mathbf{1}$ denotes a matrix of ones. Specifically, see that the kernel-weighted residuals, $K \circ \hat{e}\hat{e}'$, are

$$\begin{bmatrix} k_{11}\hat{e}_1\hat{e}_1 & k_{12}\hat{e}_1\hat{e}_2 & \dots & k_{1N}\hat{e}_1\hat{e}_N \\ k_{12}\hat{e}_1\hat{e}_2 & k_{22}\hat{e}_2\hat{e}_2 & & \\ \vdots & & \ddots & \\ k_{1N}\hat{e}_1\hat{e}_N & k_{2N}\hat{e}_2\hat{e}_N & \dots & k_{NN}\hat{e}_N\hat{e}_N \end{bmatrix} = \begin{bmatrix} k_{11} & k_{12} & \dots & k_{1N} \\ k_{12} & k_{22} & & \\ \vdots & & \ddots & \\ k_{1N} & k_{2N} & \dots & k_{NN} \end{bmatrix} \circ \begin{bmatrix} \hat{e}_1 & \hat{e}_1 & \dots & \hat{e}_1 \\ \hat{e}_2 & \hat{e}_2 & & \\ \vdots & & \ddots & \\ \hat{e}_N & \hat{e}_N & \dots & \hat{e}_N \end{bmatrix} \circ \begin{bmatrix} \hat{e}_1 & \hat{e}_2 & \dots & \hat{e}_N \\ \hat{e}_1 & \hat{e}_2 & & \\ \vdots & & \ddots & \\ \hat{e}_1 & \hat{e}_2 & \dots & \hat{e}_N \end{bmatrix}$$

By reformulating the algebra in terms of Hadamard products of sparse matrices, thousands or millions of operations are computed, rather than billions or trillions. This is because once K is separated from the residuals, it only needs to be computed once for all time periods, multiplying the outer product of the residuals by using techniques that exploit its sparseness. While I use this procedure for space-time correlations, it is also practical in other applications (for example, with a binary $K-$matrix that indicates clusters).

[1] Note that either Σ_i or Σ_t are specified with a 0 diagonal as to not double count the variance for cell g in both matrices.