We seek a computationally efficient way to construct the inverse-covariance (i.e., precision) matrix that results in a two-dimensional process with that is indexed by age and year , where the value has correlations along a year, age, and cohort axis. This inverse-covariance would then be used to specify a hyperdistribution for the process:

where is a vector that follows a multivariate normal distribution with a mean vector of 0s, and is the inverse of the precision matrix (i.e., covariance). Evaluating the multivariate normal probability density function requires computing the quadratic term precision , such that it is computationally efficient to construct directly.

To do so, we take inspiration from Simultaneous Autoregressive (SAR) processes in spatial statistics (Ver Hoef et al. 2018). Specifically, we construct a square matrix that represents the partial effect of on preceding ages and/or years (Fig. 1). The square matrix is analogous to the spatial weights matrix described in the SAR literature. Note that ages and/or years do not depend on themselves and thus, matrix has zeros on the diagonals. Furthermore, matrix **B** does not need to be symmetrical as it does not directly appear in the inverse of the covariance matrix (Eq. 3). In a simplified example of matrix **Y** indexed with ages and (e.g., where the second element of ***y*** corresponds to age 2 and year 1), we then construct matrix **B** as:

where is autocorrelation among ages in a given year, is autocorrelation among years for a given age, and is autocorrelation along a cohort. We also include a simple R-script demonstrating this construction in Appendix A. We then construct the precision matrix as:

where **I**is an identity matrix, and is a positive diagonal matrix that determines the variance of the process. This specification of the precision matrix requires that be invertible, such that the covariance matrix is positive-definite. To illustrate the effects of Eq. 2 and 3, we construct the covariance and visualize random multivariate normal draws using different specified values for partial correlations to provide intuition on scenarios with strong age, year, or cohort effects (Fig. 2).

We note two alternative ways to specify , where both involve specifying that is a diagonal matrix. We call these the “conditional variance” and “marginal variance” forms:

1. “Conditional variance” form: In the following, we specify that , where would be an estimated parameter representing the variance for conditional upon previous ages and years. This construction then results in a heteroskedastic (and potentially nonstationary) process, i.e., where varies among ages and years, but has the benefit that there are no restrictions on partial effects .
2. “Marginal variance” form: We could instead calculate values for such that , where would be an estimated parameter representing the marginal variance for , which is stationary (i.e., the same value for all ages and years). This then implies bounds on , and requires some extra code to implement (see Appendix A for details).

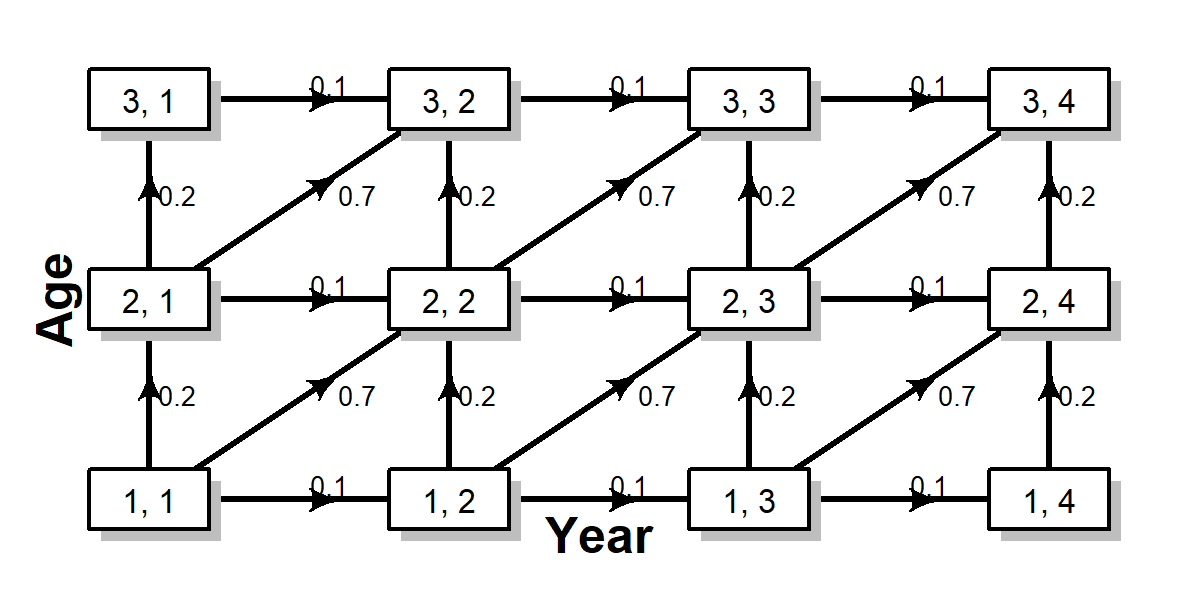


Figure 1: Diagram illustrating , i.e., the partial effect of age (y-axis), year (x-axis), and cohort (ascending diagonal) on a matrix that is indexed by age and year, in this case showing a weak partial effect of year , intermediate partial effect of age , and strong partial effect of cohort

Chart, treemap chart

Description automatically generated

Figure. 2 – Examples of random values of process drawn from a multivariate normal distribution resulting from strong year (left panel), age (middle panel), or cohort (right panel) effects

Chart

Description automatically generated

Figure 3.