McGill Computational Science and Engineering Seminar

Friday November 3rd at 14:00 in Burnside Hall room 1205

Parallel linear model estimation: (restricted) maximum likelihood & the environmental sciences

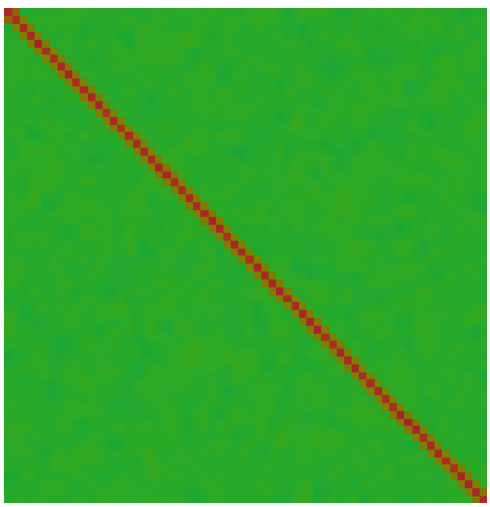
Joel Malard

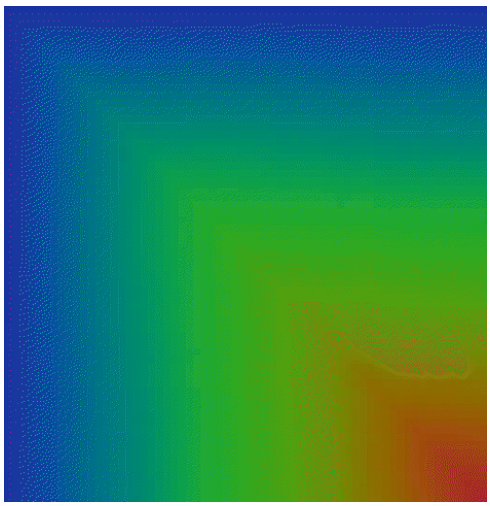
Pacific Northwest National Laboratory

This talk will demonstrate that it is possible to compute in parallel estimates of covariance matrix for linear models based on restricted maximum likelihood (REML) and to do so quite efficiently.

So, what are linear models? As you may known, there is more to a system of linear equations Ax=b than just A and b. Often the system cannot be solved exactly and an error term arises whose numerical structure is only known partially. These stochastic linear systems occur in applications such as estimating the genetic potential of cattle, financial time series analysis, the characterization of chemical mixtures and parameter estimation for subsurface flow models. Some structural information about the error that is captured in so called covariance matrices can be used to bound the uncertainty in x. Accurate estimates of covariance matrices can be very demanding and simplifying assumptions are often made in practice. Advances in computing and data collection techniques are eroding the justifications for such simplifying assumptions.

Why should you care? First, a REML code whose run time scales up to 128 processors is not a common thing yet. Most importantly, ML and REML estimations are but two examples of a wealth of stochastic computations of practical importance to environmental sciences that are becoming tractable due to advances in high-performance computing.





These are sample covariance matrices: the random vectors used for these images were generated in related ways, $u_i = (v_i + v_{i-1})/\sqrt{2}$ for the first image and $w_i = v_1 + v_2 + \ldots + v_i$ for the second image (a random walk).

Coffee and desserts are served after the seminar at 15:00 in the lounge, Burnside Hall 1024