Rolling Gaussian Process Regression

Bill Menke, June 14, 2022

Here, I describe an enhancement to real-time Gaussian Process Regression (described in the previous note), in which, after a certain number, say L, of data is reached, the earliest N_1 data are deleted before the most-recent N_1 data are added. This process can be considered rolling GPR.

The idea is that at the start of each real-time iteration, the control-to-control correlation matrix, $\mathbf{A}$, and its inverse are known and can be written in terms of submatrices

$$\mathbf{A} = \begin{bmatrix} \mathbf{X} & \mathbf{Z}^T \\ \mathbf{Z} & \mathbf{Y} \end{bmatrix} \quad \text{and} \quad \mathbf{A}^{-1} = \begin{bmatrix} \mathbf{P} & \mathbf{R}^T \\ \mathbf{R} & \mathbf{Q} \end{bmatrix} \tag{1}$$

Here, $\mathbf{X}$ and $\mathbf{P}$ are $N_1 \times N_1$, $\mathbf{Y}$ and $\mathbf{Q}$ are $K \times K$, and $\mathbf{Z}$ and $\mathbf{R}$ are $K \times N_1$, with $K + N_1 = L$. The process of removing the first N_1 data corresponds to replacing $\mathbf{A}$ with $\mathbf{Y}$, and $\mathbf{A}^{-1}$ with $\mathbf{Y}^{-1}$. To proceed, we need an efficient method for computing $\mathbf{Y}^{-1}$. We use the Woodbury formula:

$$(\mathbf{A} + \mathbf{UV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{I} + \mathbf{VA}^{-1}\mathbf{U})^{-1}\mathbf{VA}^{-1} \tag{2}$$

where $\mathbf{U}$ and $\mathbf{V}$ are chosen so that $(\mathbf{A} + \mathbf{UV})$ and its inverse are the block-diagonal matrices:

$$\mathbf{A} + \mathbf{UV} = \begin{bmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{Y} \end{bmatrix} \quad \text{and} \quad (\mathbf{A} + \mathbf{UV})^{-1} = \begin{bmatrix} \mathbf{X}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{Y}^{-1} \end{bmatrix} \tag{3}$$

Thus, $\mathbf{Y}^{-1}$ can be read from the lower-right-hand corner of $(\mathbf{A} + \mathbf{UV})^{-1}$. It can easily be verified that the correct form of $\mathbf{A} + \mathbf{UV}$ is obtained by the choices:

$$\mathbf{U}_{L,N_1} = \begin{bmatrix} \mathbf{I}_{N_1,N_1} & \mathbf{0}_{N_1,N_1} \\ \mathbf{0}_{K,N_1} & -\mathbf{Z} \end{bmatrix} \quad \text{and} \quad \mathbf{V}_{N_1,L} = \begin{bmatrix} \mathbf{0}_{N_1,N_1} & -\mathbf{Z}^T \\ \mathbf{I}_{N_1,N_1} & \mathbf{0}_{N_1,K} \end{bmatrix} \tag{4}$$

Here, subscripts are used to indicate the sizes of the zero and identity matrices. Note that the matrix, $(\mathbf{I} + \mathbf{VA}^{-1}\mathbf{U})$, which appears in the Woodbury formula is $(2N_1) \times (2N_1)$, so that its inverse requires less effort to compute than that of the $K \times K$ matrix, $\mathbf{Y}$. Some efficiency can be gained in evaluating the Woodbury formula by noting that:

$$\mathbf{I} + \mathbf{VA}^{-1}\mathbf{U} = \mathbf{I} + \begin{bmatrix} \mathbf{0} & -\mathbf{Z}^T \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P} & \mathbf{R}^T \\ \mathbf{R} & \mathbf{Q} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{Z} \end{bmatrix} = \begin{bmatrix} \left[\mathbf{I}_{N_1,N_1} - \mathbf{Z}^T\mathbf{R}\right] & \mathbf{Z}^T\mathbf{QZ} \\ \mathbf{P} & \left[\mathbf{I}_{N_1,N_1} - \mathbf{R}^T\mathbf{Z}\right] \end{bmatrix}$$

$$\mathbf{A}^{-1}\mathbf{U} = \begin{bmatrix} \mathbf{P} & \mathbf{R}^T \\ \mathbf{R} & \mathbf{Q} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{Z} \end{bmatrix} = \begin{bmatrix} \mathbf{P} & -\mathbf{R}^T\mathbf{Z} \\ \mathbf{R} & -\mathbf{QZ} \end{bmatrix}$$

$$\mathbf{VA}^{-1} = \begin{bmatrix} \mathbf{0} & -\mathbf{Z}^T \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P} & \mathbf{R}^T \\ \mathbf{R} & \mathbf{Q} \end{bmatrix} = \begin{bmatrix} -\mathbf{Z}^T\mathbf{R} & -\mathbf{Z}^T\mathbf{Q} \\ \mathbf{P} & \mathbf{R}^T \end{bmatrix} \tag{5}$$

The data removal process is then:

The current size of $\mathbf{A}$ is $L \times L$; $K = L - N_1$
Extract the $K \times N_1$ matrix, $\mathbf{Z}$, from the bottom-left corner of $\mathbf{A}$
Create the $L \times N_1$ matrix, $\mathbf{U}$, and the $N_1 \times L$ matrix, $\mathbf{V}$
Use the Woodbury formula to calculate the $L \times L$ matrix, $(\mathbf{A} + \mathbf{UV})^{-1}$ as in Eqn. (2)
Extract the $K \times K$ matrix, $\mathbf{Y}^{-1}$, from the bottom-right corner of $(\mathbf{A} + \mathbf{UV})^{-1}$
$\mathbf{A}$ becomes $\mathbf{Z}$
$\mathbf{A}^{-1}$ becomes $\mathbf{Y}^{-1}$
Delete the first N_1 elements of $\Delta\mathbf{d_1}$
Delete the first N_1 columns of $\mathbf{C}_m^{(tc1)}$
Recalculate $\hat{\mathbf{u}}_\mathbf{1} = \mathbf{A}^{-1}\Delta\mathbf{d_1}$
The current size of $\mathbf{A}$ is $(L - N_1) \times (L - N_1)$

Example: The goal of this numerical experiment is to reconstruct a two-dimensional field, $m(x, y)$, on the interval $0 \le x \le 1$, $0 \le y \le 1$, evaluated on a 30×30 grid of uniformly-spaced target points. At each time step, a total of 5 data are collected, drawn at randomly chosen points from the true function $m(x, y) = \sin(2\pi x)\sin(2\pi y)$ and with variance $\sigma_d^2 = 0.01$ The field is assumed to have the Gaussian autocovariace, $C(x, x', y, y') = \exp(-½r^2/s^2)$, with $r^2 = (x - x')^2 + (y - y')^2$ and scale length $s = 0.22$. The rolling process begins at time step 10. The reconstruction systematically improves with time, up to time 10 as new data are obtained and then stabilizes at the rolling process begins (Fig. 1).

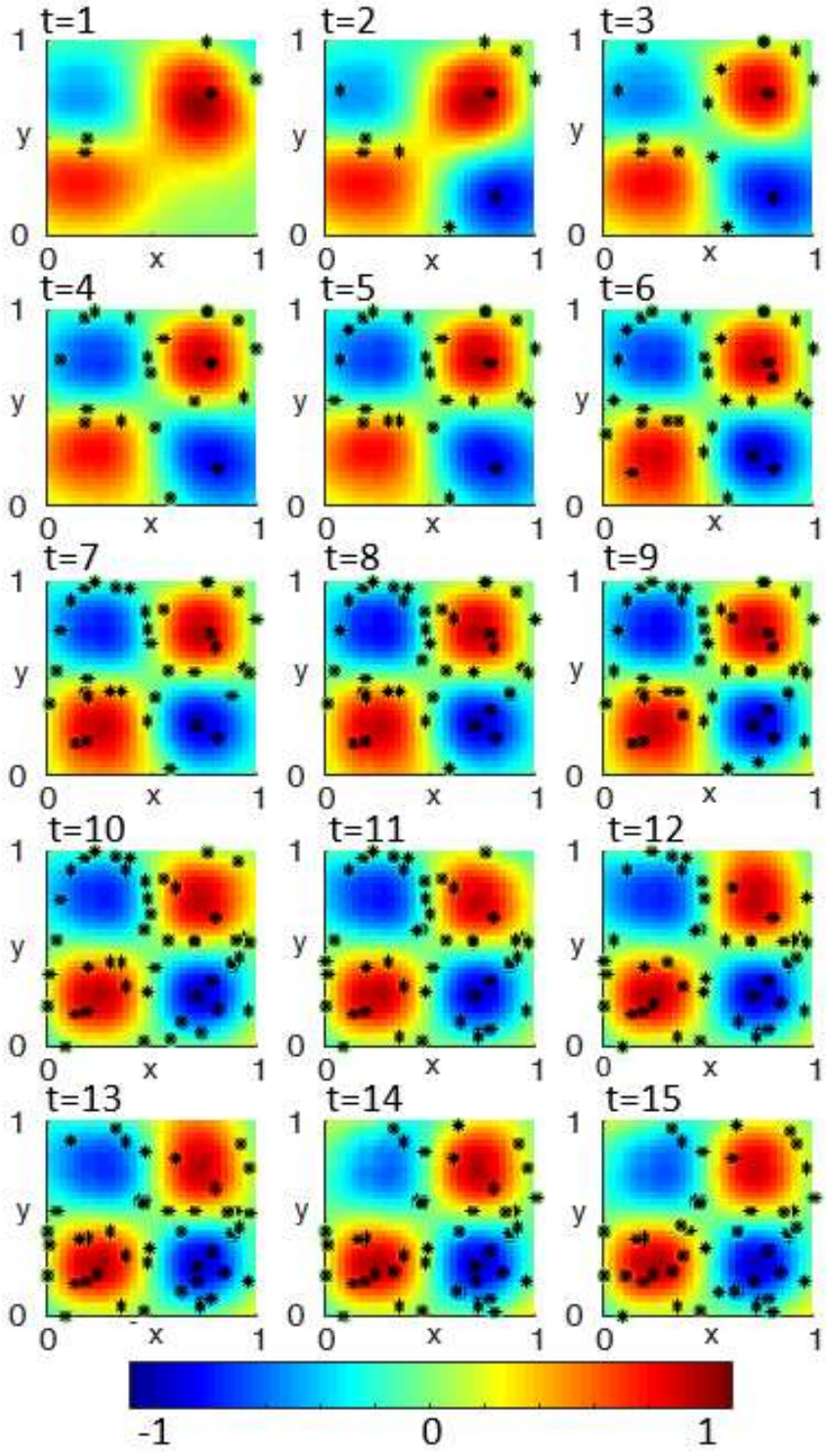

Figure 1. The sinusoidal field, $m(x, y)$, reconstructed at a sequence of 15 time steps. Five data are added in each time step; the rolling process begins at time step 10.