Circulant Embeddings for Toeplitz Covariance Matrices

James Nichols
University of New South Wales
james.nichols@student.unsw.edu.au

16 August 2010

- Guassian random field generation (usually on points $\{x_k\}_{k=1 \dots N}$), e.g. for solving a PDE with random coefficients

- Given a covariance kernel $r(|x-y|)$, e.g. $r(|x-y|) = \sigma^2 e^{-|x-y|/\lambda}$

- We have a covariance matrix R for the field on the discrete points (but there are other ways of generating the field)

- To generate points, we need to decompose R into its "square root", $R = AA^T$

- This decomposition, and the subsequent multiplications, can be very expensive, e.g. Cholesky decomposition, $O(N^3)$

- Consider a regular grid of N points on $[0, 1]$

- R will have a Toeplitz structure, i.e. constant on the diagonals

$$R = \begin{bmatrix} r(0) & r(1/N) & r(2/N) & r(3/N) & \cdots \\ r(1/N) & r(0) & r(1/N) & r(2/N) & \\ r(2/N) & r(1/N) & r(0) & r(1/N) & \\ \vdots & & & & \ddots \end{bmatrix}$$

Note that R can be characterised by its first row $\boldsymbol{r}$, where $\boldsymbol{r}_k = r(k/N)$

- We can generalise to N^s points on $[0, 1]^s$

- R will have a block-Toeplitz with Toeplitz blocks structure, i.e.

$$R = \begin{bmatrix} R_0 & R_1 & R_2 & R_3 & \cdots \\ R_1 & R_0 & R_1 & R_2 & \\ R_2 & R_1 & R_0 & R_1 & \\ \vdots & & & & \ddots \end{bmatrix}$$

Where each of the R_k is a square Toeplitz matrix

- Any Toeplitz matrix has a circulant embedding (where each row is the previous shifted one place). Consider R of size $N \times N$ embedded in our circulant C of size $d \times d$:

$$C = \begin{bmatrix} R & U \\ U^T & V \end{bmatrix}$$

  e.g. minimal embedding

$$\boldsymbol{c} = \{r_0, r_1, ..., r_{N-2}, r_{N-1}, r_{N-2}, ..., r_2, r_1\}$$

  (again a circulant can be charaterised by its first row)

- Eigenvectors of C are the columns of the DFT matrix, and C can be decomposed as $C = F^T \Lambda F$

- Can use the FFT for finding eigenvalues Λ and subsequent multiplications $\boldsymbol{z} = F\Lambda^{1/2}\boldsymbol{x}$. Cheap! $O(d \log d)$.

- But how does that help us?

$$\mathbb{E}(g(\boldsymbol{z})) = \int_{\mathbb{R}^N} g(\boldsymbol{z}) \frac{\exp(\frac{1}{2}\boldsymbol{z}^T R^{-1}\boldsymbol{z})}{\sqrt{(2\pi)^N(\det R)^{1/2}}}\,\mathrm{d}\boldsymbol{z}$$

$$= \int_{\mathbb{R}^d} g(\boldsymbol{y}_{[1\ldots N]}) \frac{\exp(\frac{1}{2}\boldsymbol{y}^T C^{-1}\boldsymbol{y})}{\sqrt{(2\pi)^N(\det C)^{1/2}}}\,\mathrm{d}\boldsymbol{y}$$

- So expectation stays the same with extra variables and extended covariance matrix

- For QMC integration we can take this to the unit cube, with $C = SS^T$

$$= \int_{[0,1]^d} g(S\Phi_d(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}$$

($\Phi_d(\boldsymbol{x})$ is the d-dimensional inverse normal)

- But circulant C is not guaranteed to be positive definite.

- Can solve by padding:

$$\boldsymbol{c} = \{r_0, ..., r_{N-1}, p_1, p_2, ..., p_k, ..., p_1, r_{N-1}, ..., r_1\}$$

- We can generate the padding in a number of ways:

  - By extending using $r(x)$ itself.

  - By extending using maximum entropy methods if we only have the covariance matrix.

- If we have $r(x)$, can simply go past 1 and reflect to obtain $c(x)$.

$$c(x) = \begin{cases} r(x) & 0 \leq x < \frac{N+k}{N} \\ r(2\frac{N+k}{N} - x) & \frac{N+k}{N} \leq x < 2\frac{N+k}{N} \end{cases}$$

- Then $\boldsymbol{c}_k = c(k/N)$ for $k = 1, ..., 2(N+k)-1$

$$\boldsymbol{c} = \{r(0), ..., r(\frac{N-1}{N}), r(\frac{N}{N}), ..., r(\frac{N+k-1}{N}), \\ r(\frac{N+k}{N}), r(\frac{N+k-1}{N}), ..., r(\frac{1}{N})\}$$

- Does this guarantee positive definiteness?

Theorem 1 *(Dietrich and Newsam, '97) Let $r(x)$ be a non-negative definite and symmetric function, with*

$$\tilde{s}_N(\omega) = r(0) + 2 \sum_{k=1}^{\infty} r(k/N) \cos(2\pi k\omega)$$

strictly positive. Then for every N there exists a positive integer M such that the vector $\boldsymbol{r}$ with entries $\boldsymbol{r}_k = r(k/N)$, $k = 1, ..., M$ has a non-negative definite minimal embedding.

- From the proof can show for $r(x) = e^{-|x|/\lambda}$, $M > O(N \log 2N)$ will ensure a non-negative definite embedding.

- Or, we also have

  Theorem 2 *If* $\boldsymbol{r} = \{r_0, ..., r_{N-1}\}$ *is a convex, decreasing and non-negative sequence, then* $\boldsymbol{r}$ *has a non-negative definite minimal embedding.*

- This holds for our typical choice of covariance $r(|x-y|) = \sigma^2 e^{-|x-y|/\lambda}$.

- Doesn't help for dimensions greater than 1.

- Finally, we can construct the padding using maximum entropy extensions, and through an algorithm of Dembo et al. '94, we have that.

  Theorem 3 *If* $\boldsymbol{r} = \{r_0, ..., r_{N-1}\}$ *is positive definite, then* R *has a circulant embedding of size* $2M \times 2M$ *where* $M \geq O(\kappa(R)^{1/2} N^{5/4})$.

  ($\kappa(R)$ is the condition number of the matrix R)

- Again a 1-dimensional result.

Work to be done

- Generalise these analytic results to s dimensions

- Find general relationships between eigenpairs of R and C.