Quadratic programming in synthesis of stationary Gaussian fields

Stefanos Kechagias
University of North Carolina at Chapel Hill

Graduate Seminar, September 24, 2014
(joint work with V. Pipiras and H. Helgason)

Outline

1. Synthesis of univariate processes using standard circulant embedding
2. Embedding methods for synthesis of 2D fields
3. Optimal circulant embedding

Fast Fourier Transform (FFT)

Discrete Fourier Transform (DFT) of N complex numbers $x_0, \ldots, x_{N-1}$

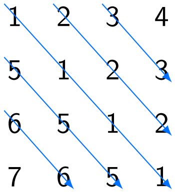
$$X_k := \sum_{n=0}^{N-1} x_n e^{-i2\pi k \frac{n}{N}}, \quad k = 0, \ldots, N-1. \tag{1}$$

- C.F.Gauss 1805, Joseph Fourier 1822, Cooley and Tukey 1965
- Direct calculation complexity is $O(N^2)$
- Complexity of FFT is $O(N \log N)$

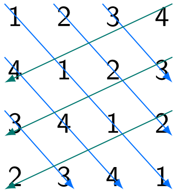
"The most important numerical algorithm of our lifetime"

Gilbert Strang

Some Linear Algebra



Toeplitz



Circulant

- Toeplitz: All diagonals remain constant
- Circulant: Each row is the preceding row shifted one entry to the right
- DFT basis diagonalizes circulant matrices. FFT can be used for efficient computation of the eigenvalues

One dimensional case

Consider a zero mean stationary Gaussian time series $\{X_n\}_{n\in\mathbb{Z}}$ with autocovariance function

$$r(n) = EX_0X_n.$$

Question: How does one generate $X := (X_0, \ldots, X_{N-1})'$, given the autocovariance function of X?

Denote the covariance matrix of X by

$$\Sigma = EXX' = \begin{pmatrix} r(0) & r(1) & r(2) & \ldots & r(N-1) \\ r(1) & r(0) & r(1) & \ldots & r(N-2) \\ r(2) & r(1) & r(0) & \ldots & r(N-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r(N-1) & r(N-2) & r(N-3) & \ldots & r(0) \end{pmatrix}$$

Cholesky

Since Σ is non-negative definite, using Cholesky decomposition we can factorize

$$\Sigma = \Sigma^{1/2}\Sigma^{1/2}$$

and set

$$X = \Sigma^{1/2} Z$$

where Z is a $N(0, I_N)$ vector.

Problem: The complexity of this method is $O(N^3)$, and the approach is practical only for moderate sample sizes N. What about larger N?

Univariate Standard Circulant Matrix Embedding (SCE)

A Circulant Matrix Embedding of Σ is a circulant matrix

$$\widetilde{\Sigma} = \begin{pmatrix} r(0) & r(1) & \dots & r(N-1) & r(N-2) & \dots & r(1) \\ r(1) & r(0) & \dots & r(N-2) & r(N-1) & \dots & r(2) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ r(N-1) & r(N-2) & \dots & r(0) & r(1) & \dots & r(N-2) \\ r(N-2) & r(N-1) & \dots & r(1) & r(0) & \dots & r(N-3) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ r(1) & r(2) & \dots & r(N-2) & r(N-3) & \dots & r(0) \end{pmatrix}$$

of dimension $M \times M$ with embedding size $M = 2N - 2$. Note that $\widetilde{\Sigma}$ contains the covariance matrix Σ . Let also

$$\widetilde{r}(n) = (r(0), r(1), \dots, r(N-1), r(N-2), \dots, r(1)) .$$

Univariate SCE

The discrete Fourier basis diagonalizes circulant matrices:

$$\widetilde{\Sigma} = F^* \Lambda F.$$

- jth column of F^* is $(1, e^{-i\frac{2\pi j}{M}}, \ldots, e^{-i\frac{2\pi j(M-1)}{M}})/\sqrt{M}$
- $\Lambda = \text{diag}(\lambda_0, \ldots, \lambda_{M-1})$, where the real eigenvalues of $\widetilde{\Sigma}$ are

$$\lambda_m = \sum_{j=0}^{M-1} \widetilde{r}(j) e^{-i\frac{2\pi j m}{M}}$$

and can be computed rapidly using FFT (complexity $O(M \log M)$).

Condition ND: The eigenvalues λ_m, $m = 0, \ldots, M-1$, are non-negative.

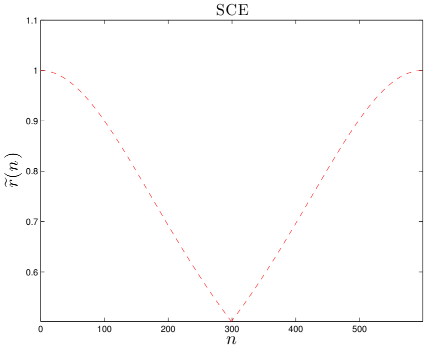
If ND holds then both

$$\widetilde{X} = \mathcal{R}(F^* \Lambda^{1/2} Z), \quad \text{and} \quad \widetilde{X} = \mathcal{I}(F^* \Lambda^{1/2} Z)$$

have covariance matrix $\widetilde{\Sigma}$.

On ND Condition

The circulant matrix embedding for the covariance structure below has 298 negative eigenvalues.



ND holds if $\{r(n)\}_{0\leq n\leq N-1}$ satisfies

- convex, decreasing, nonnegative or
- $r(k) \leq 0, \; k = 1, \ldots, N-1$ or
- N is large enough

If $\widetilde{\Sigma}$ is not smooth at the boundary of periodization, ND is likely to fail

Standard SCE for 2D fields

Consider a zero mean stationary Gaussian field $\{X_n, n \in \mathbb{Z}^2\}$ with autocovariance function

$$r(n) = r(n_1, n_2) = EX_0X_n, \quad n \in \mathbb{Z}^2.$$

Goal: Generate the field X on the square grid

$$G(N) = \{n \in \mathbb{Z}^2 : 0 \leq n_1, n_2 \leq N - 1\}.$$

given $r(n)$ in $\widetilde{G}(N) = \{n \in \mathbb{Z}^2 : -N + 1 \leq n_1, n_2 \leq N - 1\}$

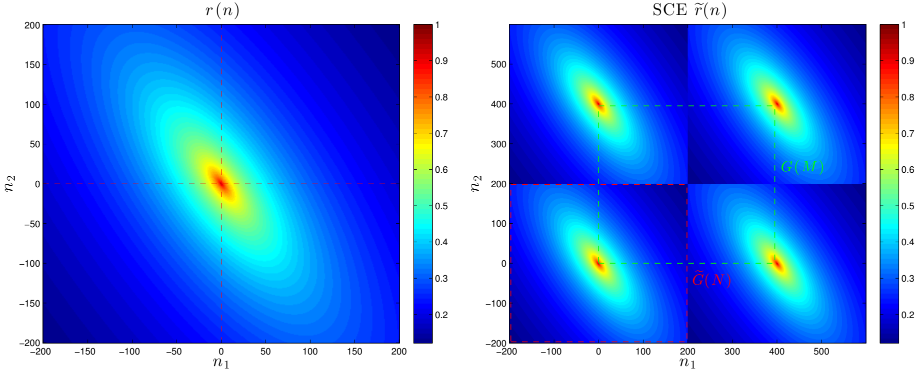
Question: What is the covariance embedding $\widetilde{r}(n)$?

Example: Powered exponential covariance function

$$r(n) = e^{-(0.01||n||_W)^{0.5}}, \quad ||t||_W := \sqrt{t'Wt}, \quad W = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

with $N = 200$ and $M = 2N - 1 = 399$.

Standard SCE for 2D fields



Extension scheme: $\widetilde{r}(n) = r(n)$, $n \in \widetilde{G}(N)$, and M-periodic

Standard SCE for 2D fields

The eigenvalues of the covariance embedding are given by the 2D DFT.

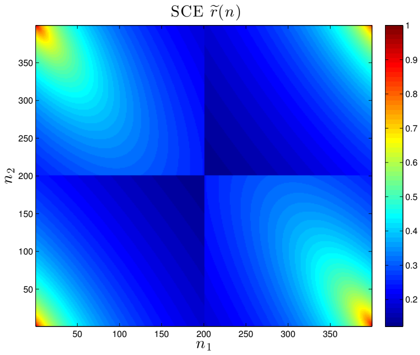
$$\lambda_k(\widetilde{r}(n)) = \sum_{n \in G(M)} \widetilde{r}(n) e^{-i2\pi k \cdot (n/M)}, \quad k \in G(M)$$

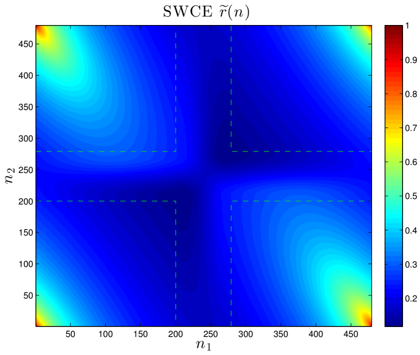
- If ND holds we can construct X as in the univariate case. Condition ND seems to fail quite often in 2D. Embedding matrix not smooth at the boundary of periodization ?

Solutions:

- Increase N to some $\widetilde{N}$ and take $M = 2\widetilde{N} - 1$. It is convenient to think of $\widetilde{G}(\widetilde{N}) \backslash \widetilde{G}(N)$ as the transition region. Increasing N to $\widetilde{N}$ in SCE can be thought as extending $r(n)$ over the transition region.
- Smoothing Windows Circulant Embedding (SWCE).
  - Apply a smoothing kernel over the transition region
  - Works well for several covariance structures.
  - Outperforms existing variants of SCE.

SCE and SWCE





Left: SCE $\widetilde{r}(n)$ for $N = 200$, $M = 399$ with 56000 negative eigenvalues.
Right: SWCE $\widetilde{r}(n)$ for $\widetilde{N} = 240$, $M = 479$ with 50 negative eigenvalues.

Optimal Circulant Embedding

The covariance embedding $\widetilde{r}(n)$ needs to satisfy the following

- $\widetilde{r}(n) = r(n)$ for n inside the dotted areas, $\widetilde{r}(n) = \widetilde{r}(-n)$
- Nonnegative eigenvalues

Idea: Obtain a circulant matrix embedding as the solution of a quadratic optimization problem with linear inequality constraints

$$\begin{aligned} \min_{\widetilde{r}(n)} \quad & f(\widetilde{r}) = \sum_{n \in G(M)} w(n)\,(r(n) - \widetilde{r}(n))^2\,, \\ \text{subject to} \quad & g_k(\widetilde{r}) \geq 0, \quad k \in G(M) \end{aligned}$$

Remark: We only need to focus on the left sub-grid L of $G(M)$, both for the objective unction and the constraints

$$\begin{aligned} \min_{\widetilde{r}(n)} \quad & f(\widetilde{r}) = \sum_{n \in L} w(n)\big(r(n) - \widetilde{r}(n)\big)^2, \\ \text{subject to} \quad & g_k(\widetilde{r}) \geq 0, \quad k \in L. \end{aligned} \tag{C}$$

Primal log-barrier method

- $\widetilde{r}(n)$ feasible point: $g_k(\widetilde{r}(n)) \geq 0, \ k \in L$
- optimal value of (C): $z^* = \inf\{f(\widetilde{r}) \mid g_k(\widetilde{r}) \geq 0, \ k \in L\}$
- $\widetilde{r}^*$ optimal point: $\widetilde{r}^*$ is feasible and $f(\widetilde{r}^*) = z^*$
- $\widetilde{r}(n)$ ϵ-suboptimal point: $\widetilde{r}$ is feasible and $f(\widetilde{r}) - z^* \leq \epsilon$

Part 1: Eliminate the constraints

$$\min_{\widetilde{r}} \quad f_t(\widetilde{r}) := f(\widetilde{r}) - \frac{1}{t}\sum_{k \in L} \log(g_k(\widetilde{r})) \qquad \text{(U)}$$

Fact: The solutions of the unconstrained problems (U), $\{\widetilde{r}_t^*, \ t > 0\}$ (*central path points*) approach a solution of the constrained problem (C) as t grows. In fact they are at most $m/t-$ suboptimal, i.e.

$$f(\widetilde{r}_t^*) - z^* \leq m/t,$$

where m is the number of inequality constraints.

Primal log-barrier method

Part 2: Quadratic approximation

For a given $x = x_0$ and a fixed t consider the second-order Taylor approximation $\widehat{f_t}$

$$\widehat{f_t}(x+v) = f_t(x) + \nabla f_t(x) v + \frac{1}{2} v^T \nabla^2 f_t(x) v$$

Calculate the direction v that minimizes $\widehat{f_t}$ (*Newton step*)

$$\min_v \quad \nabla f_t(x) v + \tfrac{1}{2} v^T \nabla^2 f_t(x) v \tag{N}$$

Fact 1: Multiple *Newton steps* yield a sequence of points $x_k = x_{k-1} + v$ that converges to the minimizer of f_t

Fact 2: We only need to take 1 Newton step!

Primal log-barrier method

Part 3: Conjugate gradient algorithm. First order conditions of (N)

$$Hv = b, \quad H = \nabla^2 f_t(x), \quad b = -\nabla f_t$$

Given v_0; Set $\epsilon_0 = Hv_0 - b$, $p_0 = \epsilon_0$, $k = 0$;
while $\epsilon_k \neq 0$

$$\begin{aligned}
\alpha_k &\leftarrow \frac{\epsilon_k^T \epsilon_k}{p_k^T H p_k}; \\
v_{k+1} &\leftarrow v_k + \alpha_k p_k; \\
\epsilon_{k+1} &\leftarrow \epsilon_k + \alpha_k H p_k; \\
s_{k+1} &\leftarrow \frac{\epsilon_{k+1}^T \epsilon_{k+1}}{\epsilon_k^T \epsilon_k}; \\
p_{k+1} &\leftarrow -\epsilon_{k+1} + s_{k+1} p_k; \\
k &\leftarrow k + 1;
\end{aligned}$$

end (**while**)

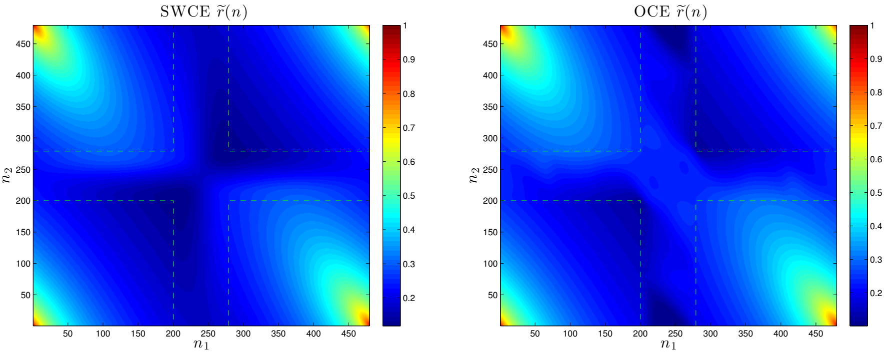
Primal log-barrier method

- Direct computation Hp_k is of order $O(N^4)$
- Using 2D FFT the order drops to $O(N^2 \log N)$
- ϵ_k doesn't need to be taken very small

Steps of the PLB method

1. Find an initial strictly feasible point $\widetilde{r}$. Pick $\mu > 1$, $t > 0$, $\epsilon > 0$.
2. Compute $\widetilde{r}^*_{t,a}$ by solving (N) with initial point $\widetilde{r}$.
3. Update $\widetilde{r} = \widetilde{r}^*_{t,a}$. If $m/t < \epsilon$, stop and return $\widetilde{r}$.
4. Increase t to μt and start again from Step 2.

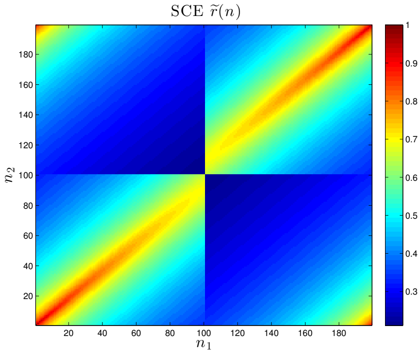
SWCE and OCE



Left: SWCE $\widetilde{r}(n)$ for $N = 200$, $\widetilde{N} = 240$, $M = 479$ with 50 negative eigenvalues.
Right: OCE $\widetilde{r}(n)$ with no negative eigenvalues and objective function value 10^{-7}.

SWCE and OCE

The magnitude of the eigenvalues of W affect the smoothness of the embedding



18000 negative eigenvalues

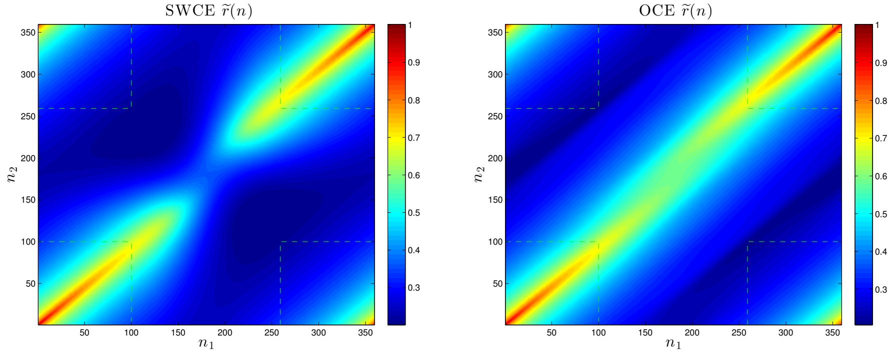
Powered exponential covariance with near singular W

$$r(n) = e^{-(0.01||n||_W)^{0.5}}$$

$$W = \begin{pmatrix} 1.6 & -1.5 \\ -1.5 & 1.4 \end{pmatrix}$$

$$\lambda_W = (0.01, 3)$$

SWCE and OCE



Left: SWCE $\widetilde{r}(n)$ for $N = 100$, $M = 359$ with 116 negative eigenvalues.
Right: OCE $\widetilde{r}(n)$ with no negative eigenvalues and objective function value 10^{-6}.

Future work

- OCE and SWCE for intrinsic Gaussian random fields
- Preconditioning in conjugate gradient algorithm
- Alternative measures - KL divergence ?
- Primal-dual path following algorithm

References

- Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix (Dietrich and Newsam), *SIAM J. Sci. Comput.*, 1997.
- Simulation of stationary Gaussian vector fields. (Chan and Wood), *Statistics and Computing*, 1999.
- Smoothing windows for the synthesis of Gaussian stationary random fields using circulant matrix embedding. (Helgason, Pipiras and Abry), J. of Comput. and Graph. Stats., 2014.
- Convex optimization and feasible circulant matrix embeddings in synthesis of stationary Gaussian fields. (Helgason, Kechagias, Pipiras), *Preprint.*

Thank you!

Remarks on OCE method

Question 1: Does the OCE method work for any covariance structure?

- As with other embedding methods, OCE's performance depends on the strength of discontinuities of the covariance embedding
- Large μ is more likely to lead to an exact embedding, which however might have some negative eigenvalues.
- Small μ ensures the eigenvalues will be nonnegative, however leading to approximate embeddings.

Question 2: How does the OCE method compares with Cholesky and SWCE in terms of speed?

- For $N > 100$ Cholesky breaks down (complexity $O(N^5)$)
- SWCE is faster at first glance. However the minimum transition region length needed is not known in advance.