Strategies for Online Inference in Dynamic Models: Estimation of a Variance Covariance-Matrix Varying Along Geodesics

Makram Talih

Department of Mathematics and Statistics
Hunter College (City University of New York)

e-mail: makram.talih@hunter.cuny.edu

August 7, 2006

Some Questions of Interest

- Covariance/Precision matrix estimation
- Matrix-valued stochastic process (eg. Wishart)
- Geometry of +ve-definite symmetric matrices $\mathcal{P} \subset \mathcal{S}$
- Online (as opposed to offline) tracking & prediction

Some Geometric Background

- Inner product on $\mathcal{S}$ is $\langle A, B\rangle = \text{tr}(AB)$, norm is $\|A\| = \sqrt{\text{tr}(A^2)}$
- $\mathcal{P}$ is a convex cone: $P + sQ \in \mathcal{P}$; $\mathcal{P}$ is open in $\mathcal{S}$
- For $P \in \mathcal{P}$, tangent space $\mathcal{T}_P$ to $\mathcal{P}$ is $\mathcal{S}$
- Define an inner product on $\mathcal{T}_P$ by
$$\langle A, B\rangle_P = \text{tr}(P^{-1}AP^{-1}B)$$
- Riemann metric on $\mathcal{P}$ is the (smooth) map $P \mapsto \langle\cdot,\cdot\rangle_P$

Geodesics

- The length of a sufficiently smooth curve $\Gamma : [a, b] \to \mathcal{P}$ is
$$\mathcal{L}(\Gamma) = \int_a^b \sqrt{\langle\dot\Gamma(s), \dot\Gamma(s)\rangle_{\Gamma(s)}}\, ds = \int_a^b \left\|\Gamma(s)^{-1}\dot\Gamma(s)\right\| ds$$
- Geodesic curves minimize length btw. $P = \Gamma(a)$ & $Q = \Gamma(b)$
- Geodesic with $\Gamma(0) = I$ & $\dot\Gamma(0) = A$ given by
$$\Gamma(s) = e^{sA}$$
- By invariance of $\mathcal{L}(\Gamma)$ to congruence, that with $\Gamma(0) = P$ is
$$\Gamma(s) = P^{1/2} e^{sP^{-1/2}AP^{-1/2}} P^{1/2}$$

Alternative DGM

$$A_{t+1} \overset{\text{IID}}{\sim} \text{Uniform on the unit sphere in } \mathbb{R}^{d(d+1)/2-1}$$

$$B_{t+1} = \|K_t^{-1} A_{t+1}\|^{-1} K_t^{-1} A_{t+1}$$

$$K_{t+1} = K_t (I - \delta B_{t+1})^{-1}, \quad 0 < \delta < 1$$

$$X_{t+1} \mid K_{t+1} \sim N_d(0, K_{t+1}^{-1})$$

Intrinsic Distance

- Intrinsic distance btw. P and Q in $\mathcal{P}$ is defined as

$$d(P,Q) = \inf\{\mathcal{L}(\Gamma) \mid \Gamma : [a,b] \to \mathcal{P}, \Gamma(a) = P, \Gamma(b) = Q\}$$

where the infimum is over sufficiently smooth curves Γ

- By the expression of the geodesic, obtain

$$d(P,Q) = \|\mathrm{Ln}(P^{-1}Q)\| = s\|P^{-1}A\|$$

- Important: $\mathcal{P}$ is geodesically complete!

Particle Filtering: Extend Step

- Maintain a (large) collection $\{\mathbb{K}_{0:t,m}\}_{m=1}^{M}$ of sample paths (called particles) leading up to time t

- Sample $\tilde{A}_m \sim$ Unif. on the unit sphere in $\mathbb{R}^{d(d+1)/2-1}$

- For each m, let

$$\tilde{B}_m = \|K_{t-1,m}^{-1} \tilde{A}_m\|^{-1} K_{t-1,m}^{-1} \tilde{A}_m$$
$$\tilde{K}_{t,m} = K_{t-1,m}(I - \delta \tilde{B}_m)^{-1}.$$

- Can use the data at $t+1$ to inform choice of candidates

- This gives an approximate sample from the posterior q_{t+1}

Data-Generating Mechanism

$$A_{t+1} \overset{\text{IID}}{\sim} \text{Unif. on the unit sphere in } \mathbb{R}^{d(d+1)/2-1}$$

$$d(K_t, K_{t+1}) \overset{\text{IID}}{\sim} \text{Exp(mean } = \delta) \quad \text{with } \delta < 1$$

$$s_{t+1} = d(K_t, K_{t+1}) / \|K_t^{-1} A_{t+1}\|$$

$$K_{t+1} = K_t^{1/2} \exp(s_{t+1} K_t^{-1/2} A_{t+1} K_t^{-1/2}) K_t^{1/2}$$

$$X_{t+1} \mid K_{t+1} \sim N_d(0, K_{t+1}^{-1}).$$

Particle Filtering: Weight, Resample & Move Step

- For each m, compute the importance weights

$$\tilde{w}_m = f_{t+1}(\boldsymbol{x}_{t+1} \mid \tilde{K}_{t+1})$$

- For each m, normalize the weights so they add up to 1:

$$w_m^* = \frac{\tilde{w}_m}{\sum_{m=1}^{M} \tilde{w}_m}$$

- If effective sample size $(\sum_{n=1}^{M} {w^*}_{t,n}^2)^{-1}$ is too small:

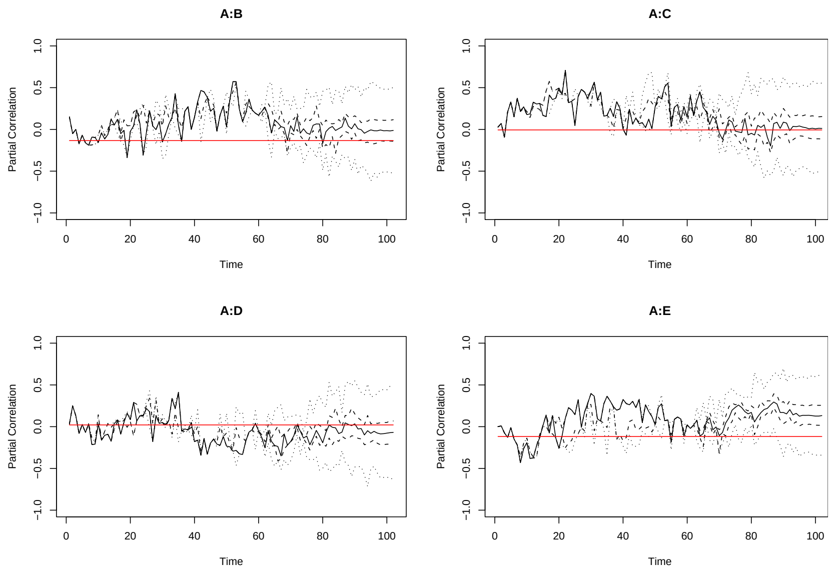
1) Sample with replacement from the set of particles according to the weights w_m^* and reset $w_m = 1/M$

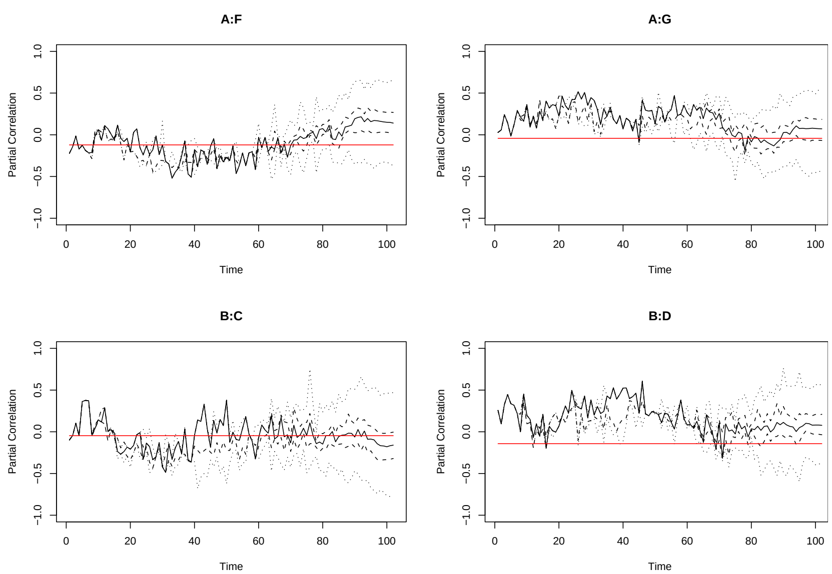
2) Apply a Gibbs/MCMC step to enhance particle diversity

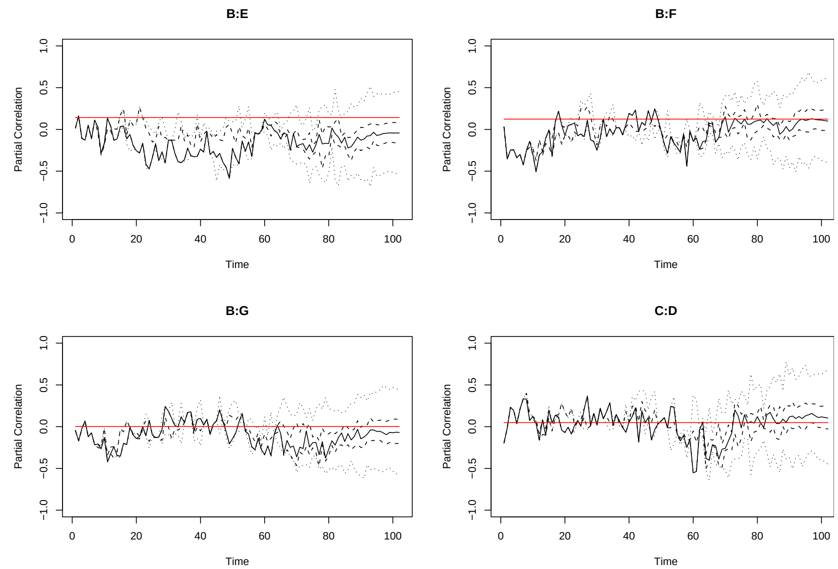
Simulation Examples...

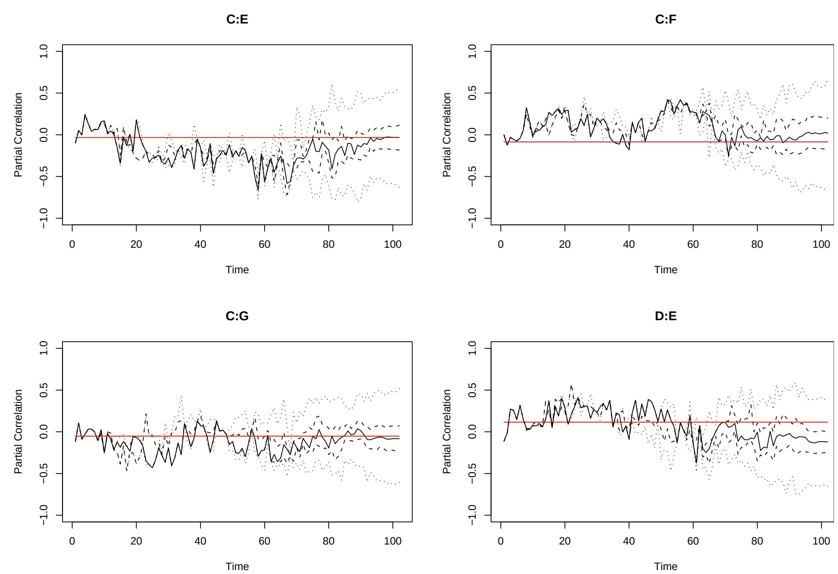
Work with standardized precision matrices
(Needs to be done after each innovation along geodesic)

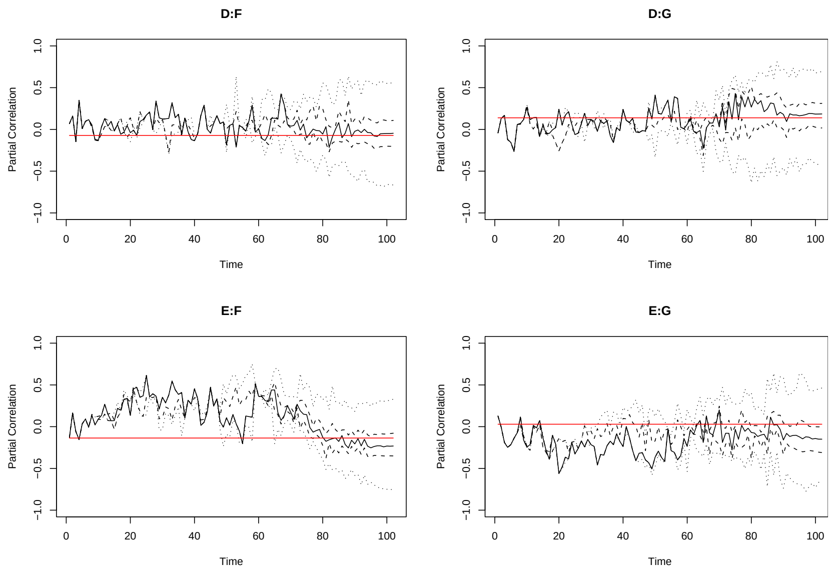
Example 1:
IID data, d=7, T=101, M=1500, ESS = 3M/4, $\delta = 0.5$, Gibbs 10 obs. deep

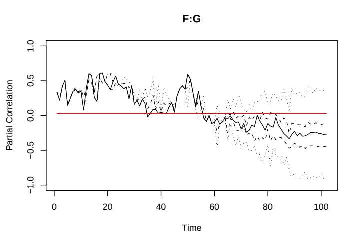





B:E
B:F
B:G
C:D
Partial Correlation
Time


C:E
C:F
C:G
D:E
Partial Correlation
Time

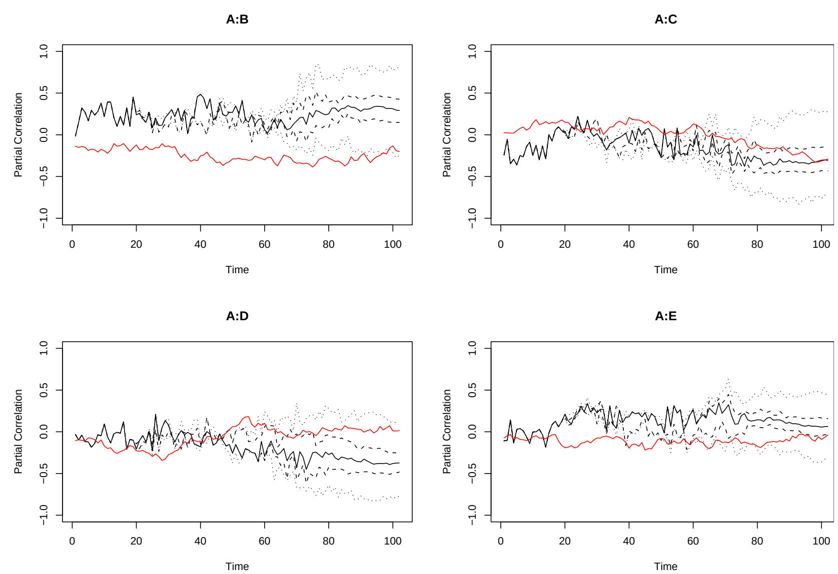

D:F
D:G
E:F
E:G
Partial Correlation
Time

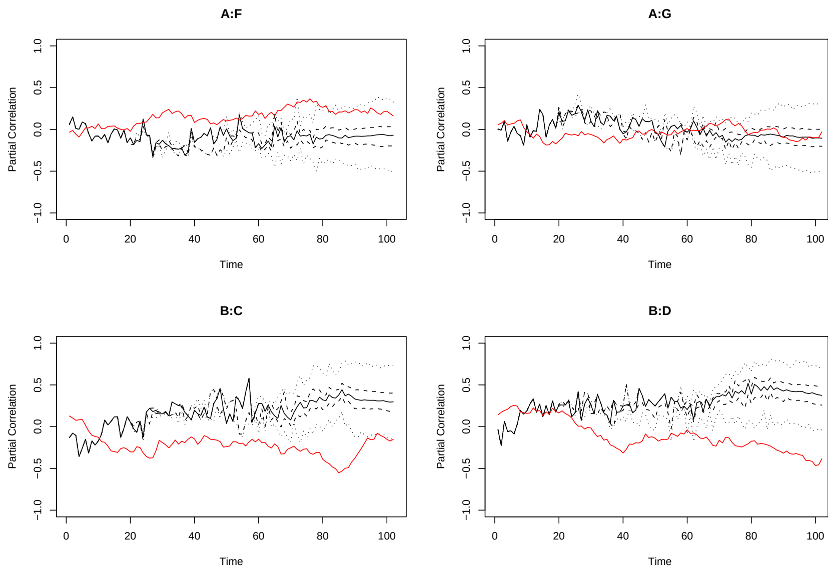

F:G
Partial Correlation
Time

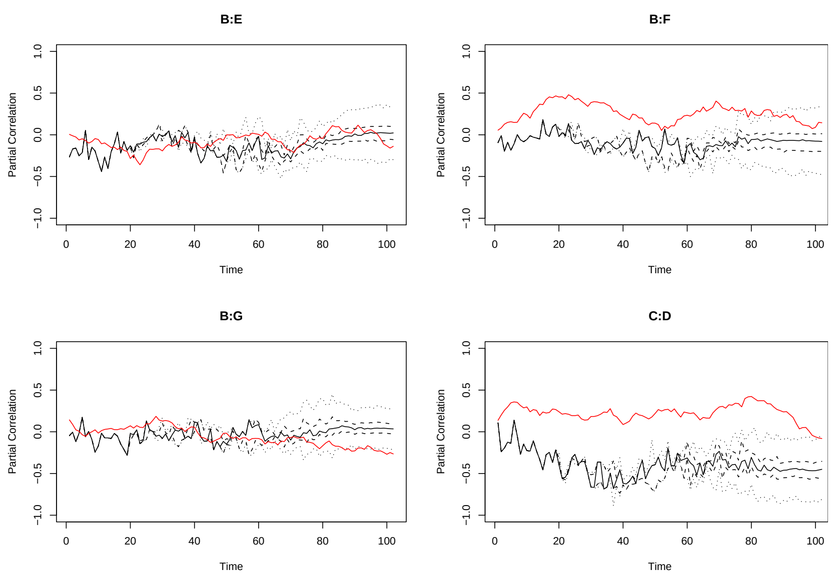
...Simulation Examples

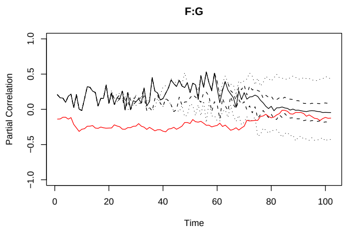
Work with standardized precision matrices
(Needs to be done after each innovation along geodesic)

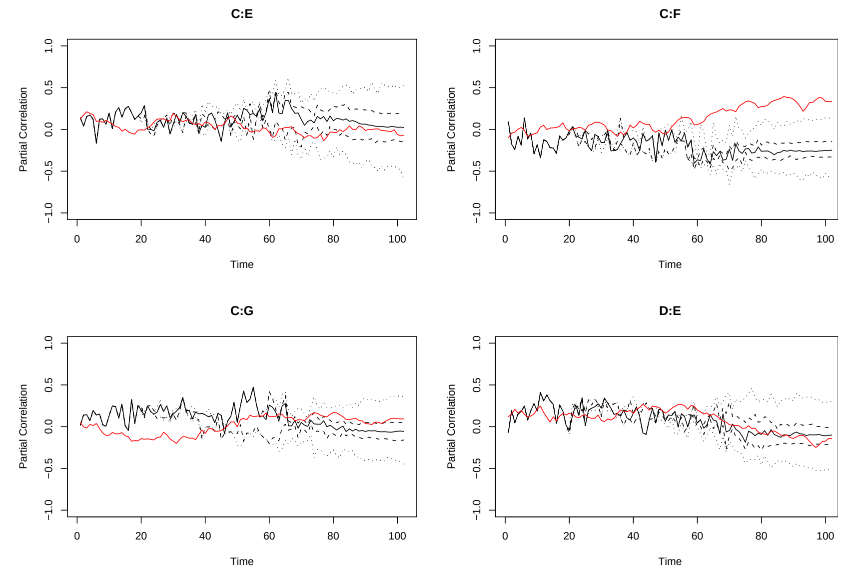
Example 2:
Time-varying data, d=7, T=101, M=1500, ESS = 3M/4, $\delta = 0.3$, Gibbs 15 obs. deep











Discussion

- Methodology for online inference in models with time-varying covariance structure

- Determine the Riemannian geometry of the cone $\mathcal{P}_G$ when a pattern of zeroes is dictated by an undirected graph G

- What is the (explicit) form of geodesics in $\mathcal{P}_G$?

- Likely problem is suspected lack of geodesic completeness in $\mathcal{P}_G$

- Design a dimension-jumping geodesic random walk?