CEE 362G Project 2

Wentao Zhang, Yimin Liu

10/18/2015

In this project, we are continuing our investigation with the linear inverse problem of image deblurring. Our focus will be on the estimation of hyperparameters, in particular two covariance weighting factors θ_1 and θ_2. Three parameter covariance types are considered, nugget covariance, linear covariance and cubic covariance. In all cases, The data covariance matrix is set to be

$$R = 10^{-6}\theta_1\delta_{ij}$$

where δ_{ij} is the Kronecker delta.

For each covariance type, two cases with different prior hyper-parameter distribution are studied. In the case 1, the probability of θ is assumed to be a constant,

$$p'(\theta_1, \theta_2) \propto 1$$

In the case 2, the probability of θ is,

$$p'(\theta_1, \theta_2) \propto \frac{1}{\theta_1\theta_2}$$

Nugget Covariance

With the nugget covariance type, covariance matrix is

$$Q = \theta_1\delta_{ij}$$

where δ_{ij} is the Kronecker delta.

First, we plot the contour map and surface map of the negative loglikelihood function with respect to θ_1 and θ_2. When selecting range of θ_1 and θ_2, we first run RML in order to find the minimum point in a very fast way. Then we calibrate out a proper range around the minimum point to best represent the function shape. Contour map and surface map for nugget covariance is shown in Figure 1.

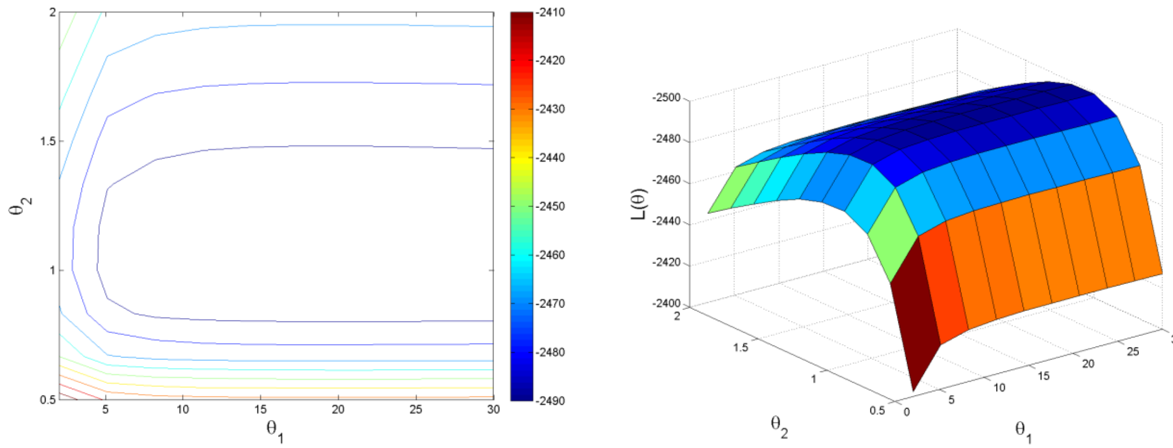

Figure 1 Contour and surface plot of negative loglikelihood function for nugget covariance. (Note z-axis direction is reverted in the surface plot)

From Figure 1, we can see curvature along the θ_2 direction is larger than curvature along the θ_1 direction. Especially when θ_1 is large, the loglikelihood function value is not very sensitive the change in θ_1. Overall the function is very smooth.

Next we solve the optimization problem using Newton-Gaussian method. For case 1, we use the *rml* function defined in *rml.m* file. For case 2, instead of modifying the rml.m file, we try to implement Newton-Gaussian by ourselves. The MAP values and approximate posterior covariance matrix for θ_1 and θ_2 are listed in Table 1 - 4.

Table 1 MAP values of θ_1 and θ_2 for nugget covariance

Case	**θ_1**	**θ_2**
1	19.0073	1.0623
2	15.4401	1.0570

Table 2 Approximate posterior covariance matrix for θ_1 and θ_2

Case 1		**Case 2**	
90.9643	-0.0016	80.4572	-0.0017
-0.0016	0.0058	-0.0017	0.0057

The best estimates, credibility intervals and conditional realizations of two cases from Nugget Covariance are shown in Figure 2~5

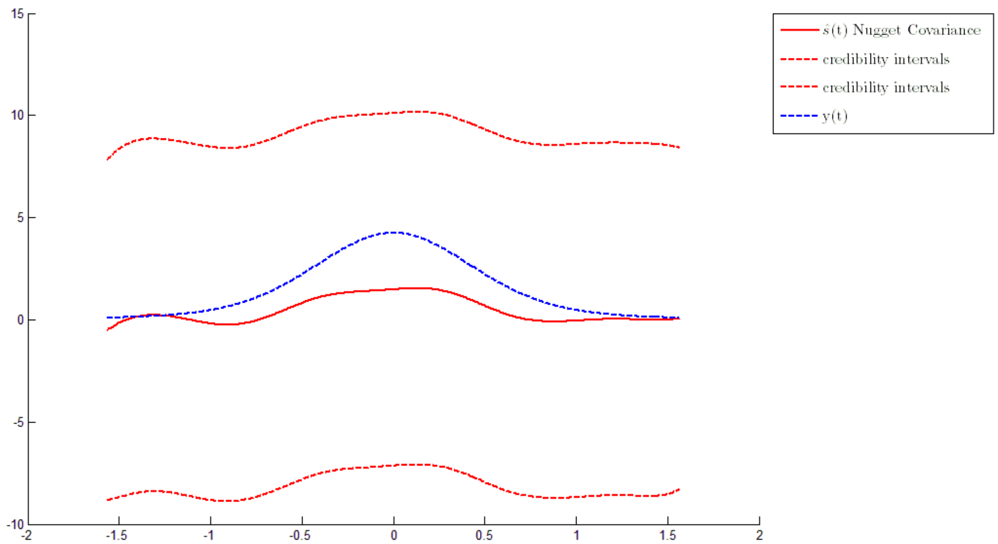

Figure 2 Best Estimates and Credibility Intervals for Nugget Covariance with Prior I

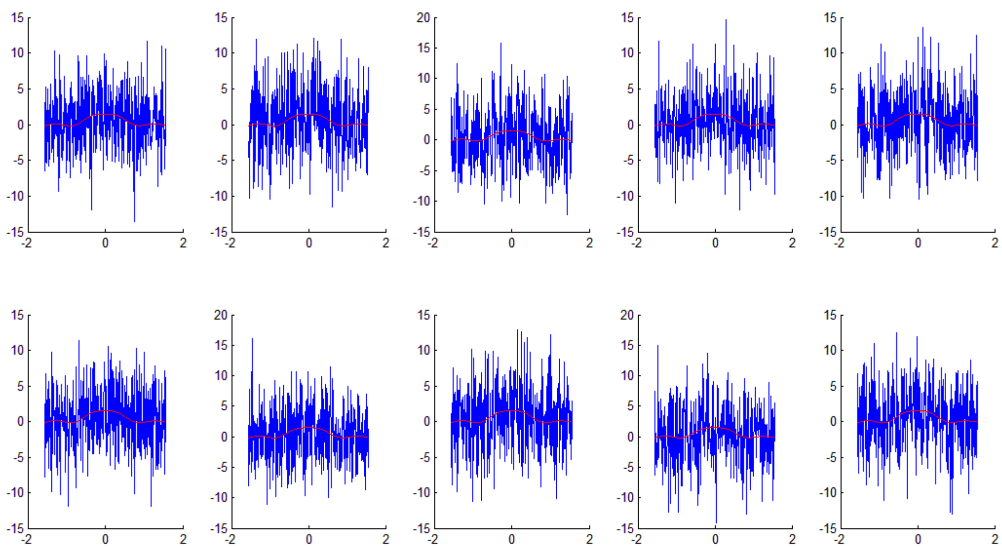

Figure 3 Realizations for Nugget Covariance with Prior I

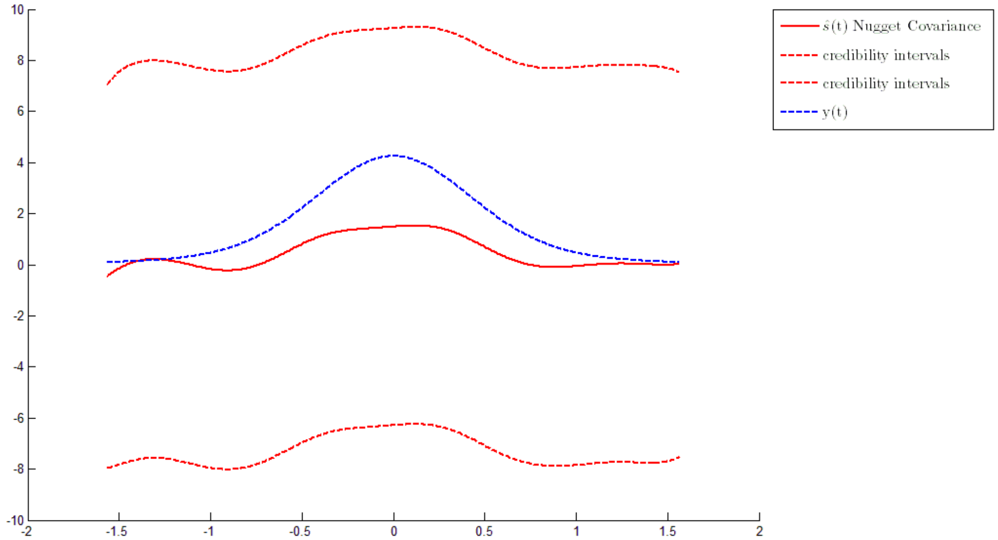

Figure 4 Best Estimates and Credibility Intervals for Nugget Covariance with Prior II

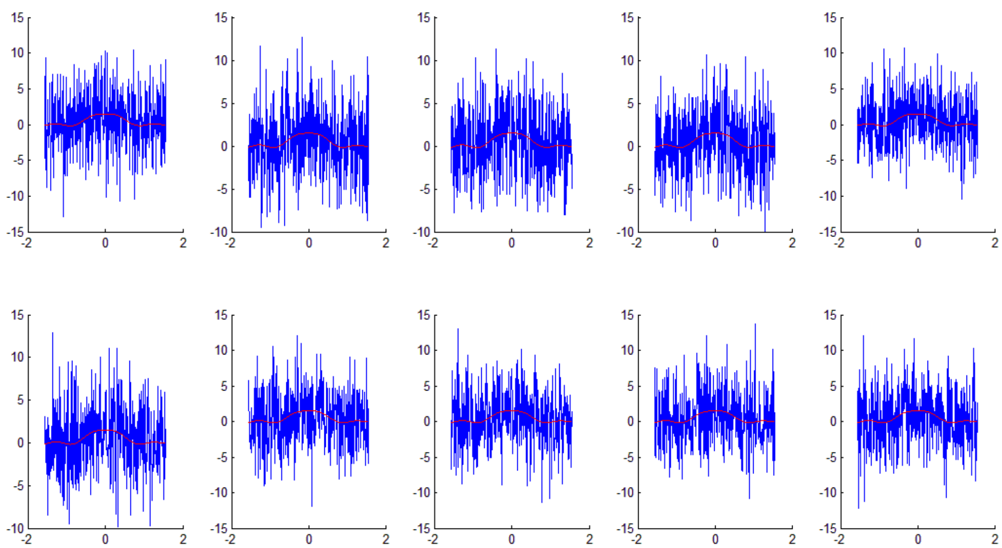

Figure 5 Realizations for Nugget Covariance with Prior II

It can be seen from the outputs that Nugget Covariance doesn't give good estimations. The posterior error covariance is quite large in both cases, deemphasizing the difference from the two priors. Also, the realizations are too noisy to recognize s.

Linear Covariance

The linear covariance matrix is as following:

$$Q = -\theta_1 ||x_i - x_j||$$

The contour and surface plot in shown in Figure 6:

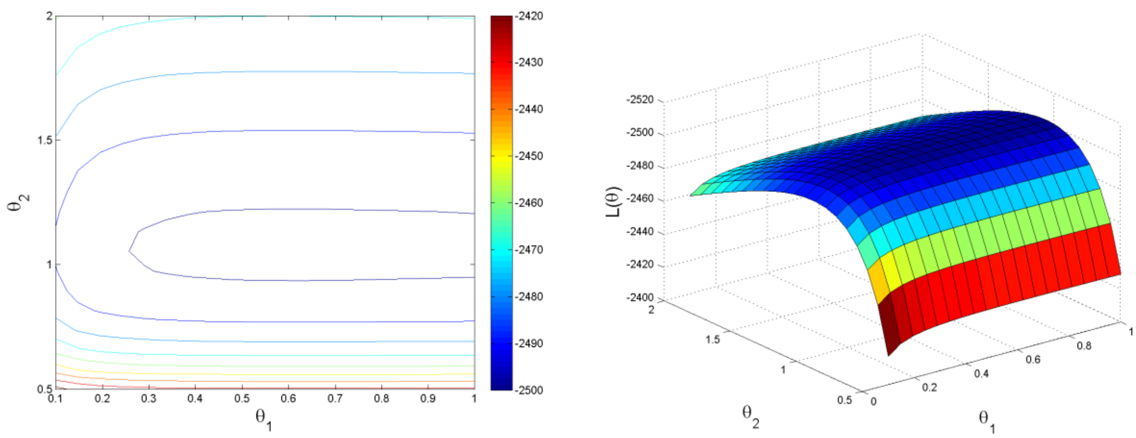

Figure 6 Contour (left) and surface plot (right) of negative loglikelihood function for linear covariance. (Note z-axis direction is reverted in the surface plot)

Table 3 MAP values of θ_1 and θ_2 for linear covariance

Case	θ_1	θ_2
1	0.6116	1.0630
2	0.4705	1.0579

Table 4 Approximate posterior covariance matrix for θ_1 and θ_2

Case 1		Case 2	
0.0986	-0.0001	0.0803	-0.0001
-0.0001	0.0058	-0.0001	0.0058

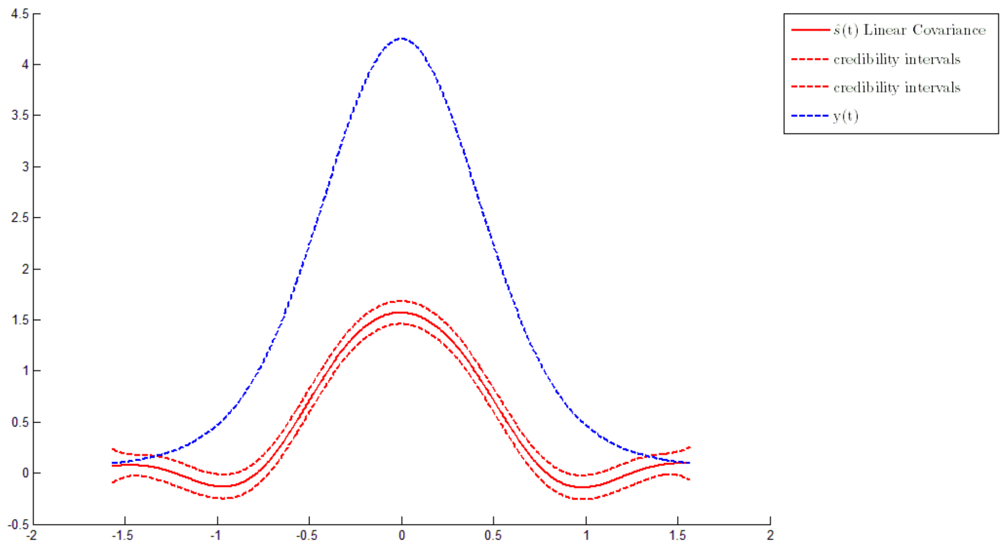

Figure 7 Best Estimates and Credibility Intervals for Linear Covariance with Prior I

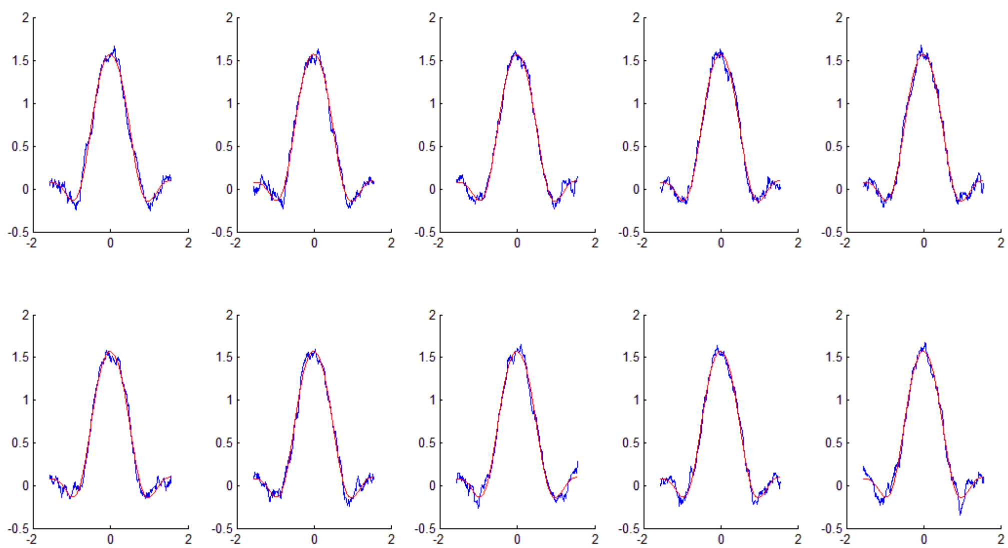

Figure 8 Realizations for Linear Covariance with Prior I

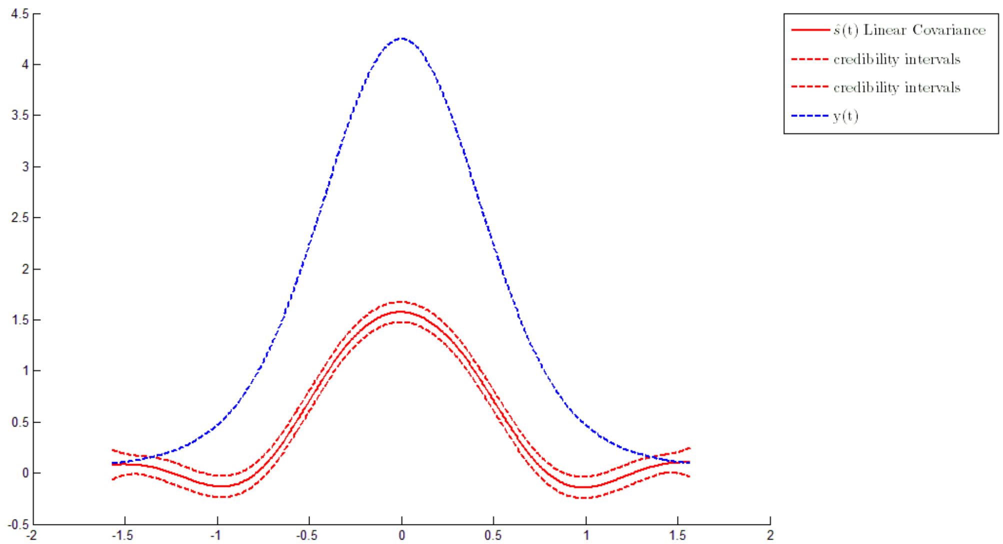

Figure 9 Best Estimates and Credibility Intervals for Linear Covariance with Prior II

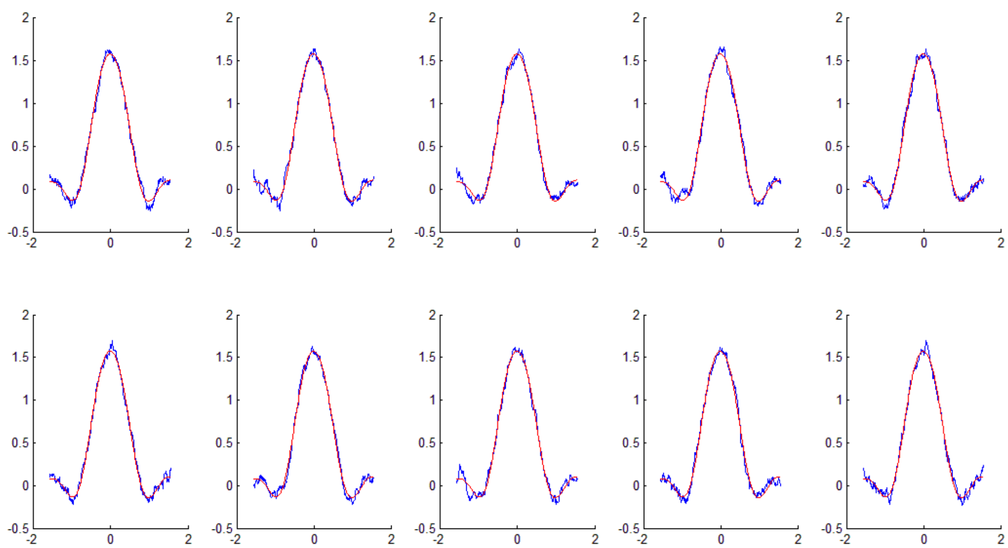

Figure 10 Realizations for Linear Covariance with Prior II

In Figure 7~10, the posterior error covariance of s from Linear Covariance estimation is much smaller than that from Nugget Covariance estimation. The smoother realizations of s indicate that the Linear estimation has a large portion in the row space of H. Again, the two types of priors don't show much difference.

Cubic Covariance

The cubic covariance matrix has the following form:

$$Q = \theta_1 ||x_i - x_j||^3$$

In this case the **X** matrix, which contains our prior knowledge of the mean of ***s*** has two columns, the first of ones and the second with coordinates x_i. The second column corresponds to a trend in the mean of **s** that is proportional to the coordinates.

The contour and surface plot in shown in Figure 11:

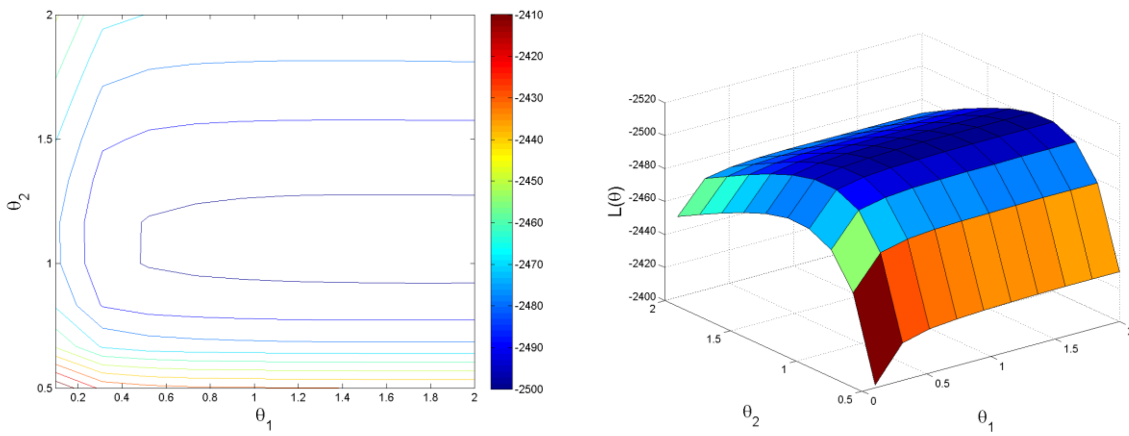

Figure 1 Contour (left) and surface plot (right) of negative loglikelihood function for cubic covariance.

Table 5 MAP values of θ_1 and θ_2 for cubic covariance

Case	θ_1	θ_2
1	1.6526	1.0664
2	1.1594	1.0631

Table 6 Approximate posterior covariance matrix for θ_1 and θ_2

Case 1		Case 2	
0.8648	-0.0004	0.6411	-0.0004
-0.0004	0.0058	-0.0004	0.0058

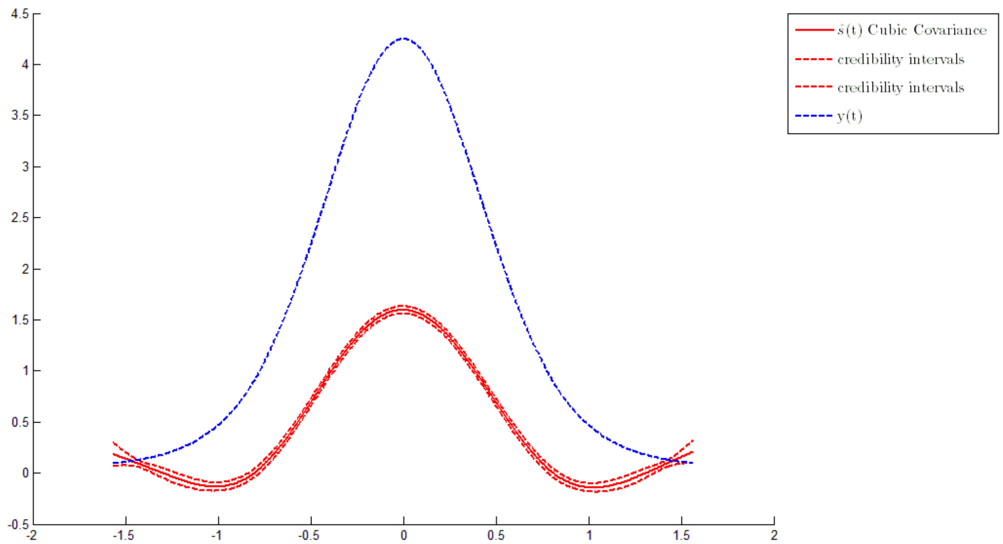

Figure 12 Best Estimates and Credibility Intervals for Cubic Covariance with Prior I

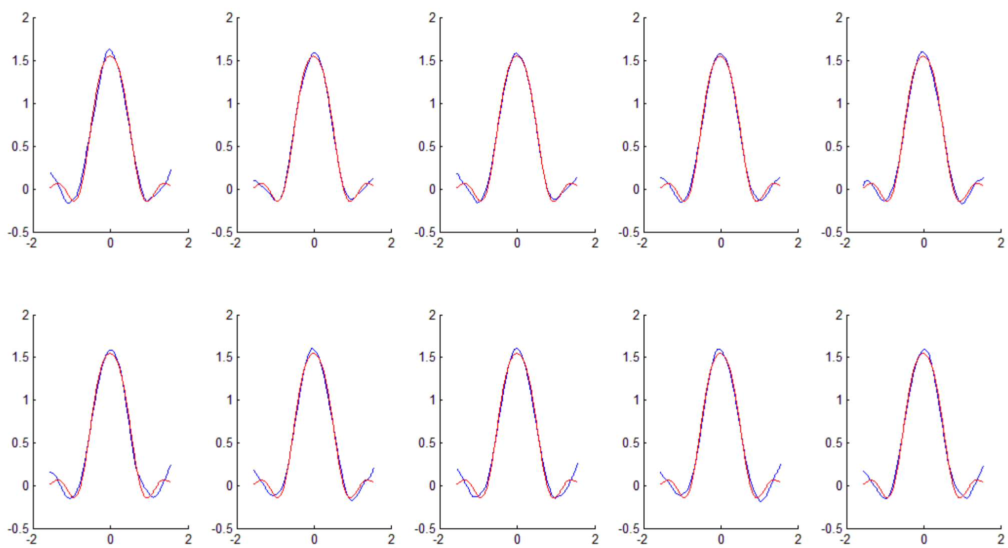

Figure 2 Realizations for Cubic Covariance with Prior I

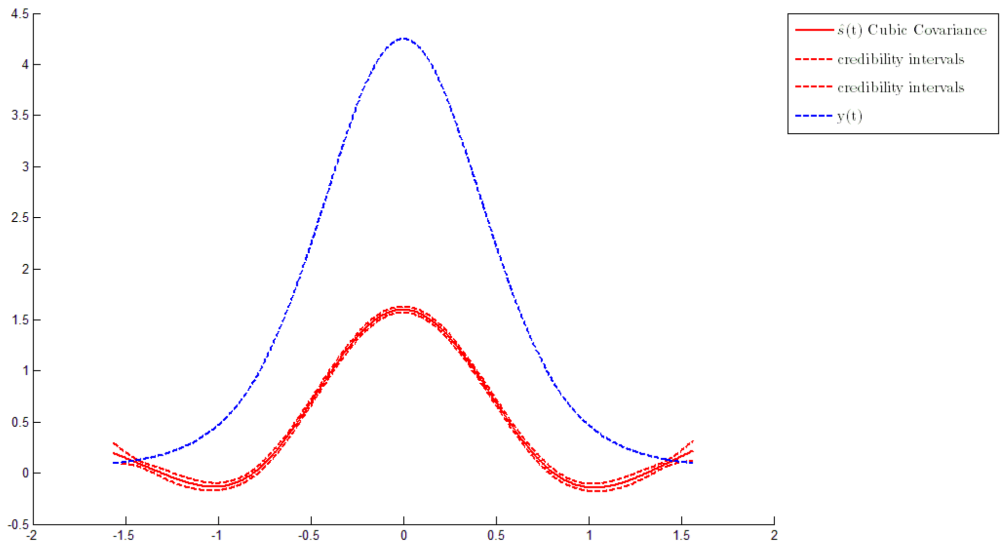

Figure 3 Best Estimates and Credibility Intervals for Cubic Covariance with Prior II

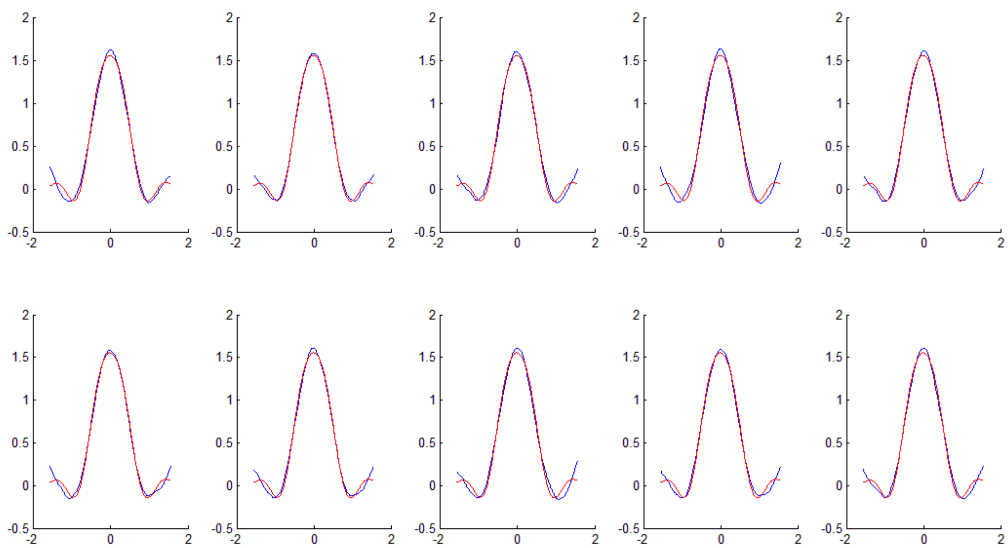

Figure 4 Realizations for Cubic Covariance with Prior II

In Figure 11~15, the uncertainty of s is further reduced by the Cubic Covariance. The extra column with coordinates in X doesn't present many impacts on the estimation results because no trend with the coordinates is observed.

Conclusion

Some important lessons we learned from this study are listed below:

(1) MAP values of θ in case 2 are consistently smaller than in case 1 with all the three covariance types. The reason is quite clear. Since the prior probability of θ is bigger when θ is smaller, the optimal point of the posterior PDF of θ will also be shifted towards smaller values.

(2) The posterior variance of θ_1 is higher than that of θ_2. We suppose this is related to the fact that the negative loglikelihood function is much more flat along θ_1 direction. Therefore, the MAP estimation of θ_1 has much higher uncertainty.

(3) Covariance types have significant impacts on both MAP values of θ and conditional realizations of s. Prior types of θ affect MAP values of θ but do not affect the estimation of s.

(4) Cubic covariance type yields smoothest results among all the three covariance types studied, followed by linear covariance type. Realizations with nugget covariance type are very noisy. In addition, since the vectors in the row space of H are very smooth, the ranking of posterior uncertainty (or the width of the credibility intervals) is the same as the ranking of smoothness.

(5) In case with cubic covariance, the prior knowledge of a trend, which is proportional to the coordinates, does not propagate into the posterior realizations. This is probably because of the credibility interval in this case is very narrow. As a result, all realizations are very close to the best estimate, which itself is very symmetric instead of following the prior trend in X.