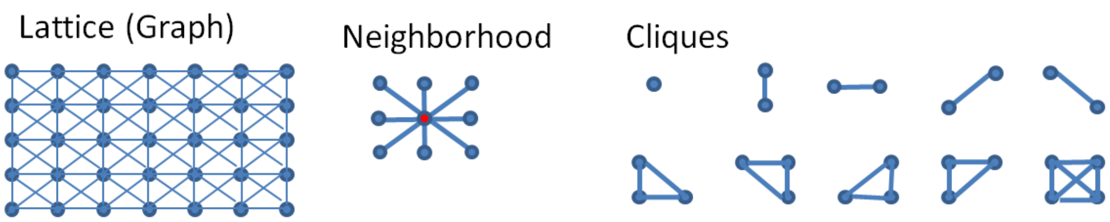
STK 9150/4150. Solution sketch.

1a) 14,15,16,24,26,34,35,36. A clique is a set of nodes in the lattice, that consist either of a single site or of a set of nodes that are all neighbors.



1b) if the subset c is not a clique then $G_c(\boldsymbol{y}_c) \equiv 0$.

If this ia s CAR model then all G's corresponding to a subset larger than 2 is zero, all G's corresponding to a subset of size 2 are constant, and all G's corresponding to a subset of size 1 are linear.

1c) Yes it is symmetric since the neighbourhood relations are symmetric.

All the eigenvalues of Q must be positive. $\frac{1}{\lambda_{\min}} < \phi < \frac{1}{\lambda_{\max}}, \quad \frac{-1}{3.9741} < \phi < \frac{1}{7.9449}$

The figure indicate edge effects. In this case it is a relatively small effect the variance increase from about 1 to 1.4 very fast. In general it is hard to balance the coefficients theoreticall, bu it is possible to define the model on a larger grid such that the effect is reduced in the region where we are interested in the results.

2a) Model 1 is the one with the best AIC and is preffered, but note that the range is very large compared to the region for which the data are collected.
Model 2 show that the second coefficient is significant thus there is evidence for a trend corresponding to the covariate x1.
The second covariate is not significant neither in model Model 3 nor Model 4 and does not bring anything new into the discussion. It is interesting to compare the range estimates in the models with and without covariate x1. The range decreases dramatically which indicates that there is a long range correlation hidden in covariate x1. Thus it is possible to argue that Model 2 should be used, but in either case it is problematic to use the models far away from the observed locations.

2b) Empirical Bayes estimate or a plugin estimator. The Empirical Bayes estimator is fast to compute and often give a satisfactory result. It does however not fully account for the uncertainties in the model. The Bayesian approach accounts for all uncertainties included in the prior distribution thus gives a more realistic estimate of the uncertainty. This approach is however more computational complex and you are required to define the prior distributions for all parameters. These can often be hard to assess. To do the Bayesian computations you can use INLA or MCMC.

2c) $E(\boldsymbol{Y}_{2|1}) = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{b} - \boldsymbol{\mu}_1)$ by linearity of expectation

$\text{Cov}(\boldsymbol{Y}_{2|1}) =$

$$\text{Cov}(\boldsymbol{Y}_2) + \text{Cov}\left(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{b} - \boldsymbol{Y}_1)\right) + \text{Cov}(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{b} - \boldsymbol{Y}_1), \boldsymbol{Y}_2) + \text{Cov}\left(\boldsymbol{Y}_2, \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{b} - \boldsymbol{Y}_1)\right)$$

$$= \boldsymbol{\Sigma}_{22} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\text{Cov}(\boldsymbol{Y}_1)\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\text{Cov}(\boldsymbol{Y}_1, \boldsymbol{Y}_2) - \text{Cov}(\boldsymbol{Y}_2, \boldsymbol{Y}_1)\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$$

$$= \boldsymbol{\Sigma}_{22} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{11}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$$

$$= \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$$

Note that this matches the expressions for conditional distribution of Y2 given Y1.

Thus by simulation from the distribution of Y and solve the equation inserting b=y1 you will obtain a correctly conditioned sample.

3a) $\boldsymbol{\mu_Y} = \boldsymbol{\mu} + \boldsymbol{M\mu_Y}, \boldsymbol{C_Y} = \boldsymbol{MC_YM^T} + \boldsymbol{C_\delta}$. Note that $\boldsymbol{C_Y} - \boldsymbol{MC_YM^T}$ must be positive definite.
Yes. In general this is not true, but the Gaussian distribution is defined by two first moments so this is sufficient.

3b) $P(\boldsymbol{Y}) = P(\boldsymbol{Y}_1)\prod_{t=2}^{T} P(\boldsymbol{Y}_t|\boldsymbol{Y}_{t-1})$.

The only terms which involve $\boldsymbol{Y}_t$ in the equation is $P(\boldsymbol{Y}_t|\boldsymbol{Y}_{t-1})$ and $P(\boldsymbol{Y}_{t+1}|\boldsymbol{Y}_t)$. Thus the distribution of $P(\boldsymbol{Y}_t|\boldsymbol{Y}_s\ s \neq t) \propto P(\boldsymbol{Y}) \propto P(\boldsymbol{Y}_t|\boldsymbol{Y}_{t-1})P(\boldsymbol{Y}_{t+1}|\boldsymbol{Y}_t)$. Thus the conditional distribution only depend on $\boldsymbol{Y}_{t-1}$ and $\boldsymbol{Y}_{t+1}$ in addition $to\ \boldsymbol{Y}_t$. Thus $P(\boldsymbol{Y}_t|\boldsymbol{Y}_s\ s \neq t) = P\left(\boldsymbol{Y}_t\middle|\boldsymbol{Y}_{t-1,}\boldsymbol{Y}_{t+1,}\right)$

log($P\left(\boldsymbol{Y}_t\middle|\boldsymbol{Y}_{t-1,}\boldsymbol{Y}_{t+1,}\right)$)=const +log($P(\boldsymbol{Y}_t|\boldsymbol{Y}_{t-1})$) +log($P(\boldsymbol{Y}_{t+1}|\boldsymbol{Y}_t)$)

$$= const_2 - 0.5 * \left((\boldsymbol{Y}_t - \boldsymbol{MY}_{t-1})^T\ \boldsymbol{C}_\delta^{-1}(\boldsymbol{Y}_t - \boldsymbol{MY}_{t-1}) + (\boldsymbol{Y}_{t+1} - \boldsymbol{MY}_t)^T\ \boldsymbol{C}_\delta^{-1}(\boldsymbol{Y}_{t+1} - \boldsymbol{MY}_t)\right)$$

$$= const_3 - 0.5 * \left(\boldsymbol{Y}_t^T\ (\ \boldsymbol{C}_\delta^{-1} + \boldsymbol{M}^T\ \boldsymbol{C}_\delta^{-1}\boldsymbol{M})\boldsymbol{Y}_t - 2\boldsymbol{Y}_t^T(\ \boldsymbol{C}_\delta^{-1}\boldsymbol{MY}_{t-1} + \boldsymbol{M}^T\ \boldsymbol{C}_\delta^{-1}\boldsymbol{Y}_{t+1}\ )\right)$$

$$= const_4 - 0.5 * (\boldsymbol{Y}_t^T\ \boldsymbol{QY}_t - 2\boldsymbol{Y}_t^T\boldsymbol{Q\mu})$$

$$\boldsymbol{Q} = (\ \boldsymbol{C}_\delta^{-1} + \boldsymbol{M}^T\ \boldsymbol{C}_\delta^{-1}\boldsymbol{M})$$

$$\boldsymbol{\mu} = \boldsymbol{Q}^{-1}(\ \boldsymbol{C}_\delta^{-1}\boldsymbol{MY}_{t-1} + \boldsymbol{M}^T\ \boldsymbol{C}_\delta^{-1}\boldsymbol{Y}_{t+1})$$

3c)

$$E\left(\begin{bmatrix} Y_{t-1} \\ Y_t \end{bmatrix}\right) = \begin{bmatrix} \mu_Y \\ \mu_Y \end{bmatrix}$$

$$\text{Cov}(\begin{bmatrix} Y_{t-1} \\ Y_t \end{bmatrix}) = \begin{bmatrix} C_Y & C_Y M^T \\ M C_Y & C_Y \end{bmatrix}$$

$$E([Y_{t-1}|Y_t]) = \mu_Y + C_Y M^T C_Y^{-1}(Y_t - \mu_Y)$$

$$\text{Cov}(Y_{t-1}|Y_t\ ) = C_Y - C_Y M^T C_Y^{-1} M C_Y$$

$$\tilde{\mu} = \mu_Y - C_Y M^T C_Y^{-1} \mu_Y$$

$$\widetilde{M} = C_Y M^T C_Y^{-1}$$

$$C_{\widetilde{\delta}} = C_Y - C_Y M^T C_Y^{-1} M C_Y$$

If there was correlation between $\widetilde{\delta_t}$ and Y_t then the variance of Y_{t-1} would not preserve stationarity. If $\widetilde{\delta_t}$ should correlate with Y_{t+i}, and not with Y_{t+i-1}, then Y_{t+i} and δ_{t+i} by construction need to be correlated, (which they are not by the argument above). thus $\widetilde{\delta_t}$ is not correlated.

3d) The Kalman filter is directional and condition to all data in the past including the current time. The Kalman smoother condition to data for past present and future of a given time, i.e.

Kalman filter : $Y_t|Z_{1:t}$ Kalman smoother: $Y_t|Z_{1:T}$.

Forward: (t,u): (1,-), (1,1)(2,1),(2,2)(3,2)(3,3)….(T,T-1),(T,T)

Backward (t,u): (T,-) (T,T)(T-1,T)(T-1,T-1),…(2,3)(2,2)(1,2)(1,1)

3e) only for STK 9150 When $\prod_{t=1}^{T} M = M^T$ (not transpose but the power of T) is small. That is when the time series is long. Since M always shrink the Covariance, this is will be the case eventually with the exception of C_δ is degenerate. By construction the forward sequence series $Y_{t+1}\ |Z_{t+1:T}$ is independent of $Y_{t-1}\ |Z_{1:t-1}$

$$Y_t \ = \ AY_{t-1} \ + BY_{t+1} + \delta$$

A, B and covariance of delta denoted C is given by c.

The distribution of the components given the data to the right and left are:

$$Y_{t-1}\ |Z_{1:t-1} \sim N(Y_{t-1|t-1}\ , P_{t-1|t-1}\ )$$

$$Y_{t+1}\ |Z_{t+1:T} \sim N(\tilde{Y}_{t+1|t+1}\ , \tilde{P}_{t+1|t+1}\ )$$

Thus:

$$\mu_{t|-t} \overset{\text{def}}{=} E\{Y_t|Z_s, s \neq t\ \} = A\ Y_{t-1|t-1}\ + B\tilde{Y}_{t+1|t+1}$$

$$C_{t|-t} \overset{\text{def}}{=} \text{Cov}\{Y_t|Z_s, s \neq t\ \} = A\ P_{t-1|t-1}\, A^T + B\tilde{P}_{t+1|t+1}\, B^T + C$$

$$E\{Y_t|Z_s, \forall s\} = \mu_{t|-t} + C_{t|-t} H_t^T \left(H_t C_{t|-t} H_t^T\right)^{-1} (z_t - H_t \mu_{t|-t})$$

$$\text{Cov}\{Y_t|Z_s, \forall\ s\ \} = \ C_{t|-t} - C_{t|-t} H_t^T \left(H_t C_{t|-t} H_t^T\right)^{-1} H_t C_{t|-t}$$