Solution 1 (to exercise 2) (a). We have

$$\mathbf{v}^T\mathbf{Z} = \mathbf{v}^T\mathbf{Y} + \mathbf{v}^T\boldsymbol{\eta}$$

Here $\mathbf{v}^T\mathbf{Y}$ is normal since $\mathbf{Y}$ is multivariate normal (from exercise 1). Further, $\mathbf{v}^T\boldsymbol{\eta}$ is a linear combination of independent normals, so is normal. Finally, from the model assumptions, $\mathbf{v}^T\mathbf{Y}$ and $\mathbf{v}^T\boldsymbol{\eta}$ are indpendent giving that $\mathbf{v}^T\mathbf{Z}$ is a linear combination of independent normals, which then is normal itself. Since this is true for all $\mathbf{v}$, we have that $\mathbf{Z}$ is normal.

(b). We have that

$$E[Z_t] = E[E[Z_t|Y_t]] = E[Y_t] = 0$$

where the last equality is from exercise 1 using that $\mu = 0$ in this case.

(c). We have that

$$\begin{aligned}\text{cov}(Z_t, Z_{t+\tau}) =& E[\text{cov}(Z_t, Z_{t+\tau})|Y_t, Y_{t+\tau}] + \text{cov}(E[Z_t|Y_t, Y_{t+\tau}], E[Z_{t+\tau}|Y_t, Y_{t+\tau}])\\ =& E[\text{cov}(\eta_t, \eta_{t+\tau})|Y_t, Y_{t+\tau}] + \text{cov}(Y_t, Y_{t+\tau})\\ =& \tau^2 I(\tau = 0) + \text{cov}(Y_t, Y_{t+\tau})\end{aligned}$$

where the first term is obtained by the properties of the η_t's.

(d). We have that

$$f(\mathbf{z}; \boldsymbol{\theta}) = f(z_1; \boldsymbol{\theta}) \prod_{t=2}^{n} f(z_t|z_1, ..., z_{t-1}; \boldsymbol{\theta})$$

Now $\mathbf{Z}$ is multivariate Gaussian. This imply that any subset of $\mathbf{Z}$ is (multivariate) Gaussian. That again imply that Z_1 is Gaussian. Further (Z_1, Z_2) is Gaussian which imply that $Z_2|Z_1$ is Gaussian and similarly $Z_t|Z_1, ..., Z_{t-1}$ is Gaussian. The result then follows.

(e). We have that

$$\hat{z}_t \equiv E[Z_t|Z_1, ..., Z_{t-1}] = E[E[Z_t|Z_1, ..., Z_{t-1}, Y_t]] = E[Y_t|Z_1, ..., Z_{t-1}] \equiv \hat{y}_{t|t-1}$$

Further,

$$\begin{aligned}S_T \equiv& \text{var}(Z_t|Z_1, ..., Z_{t-1})\\ =& \text{var}(E[Z_t|Z_1, ..., Z_{t-1}, Y_t]) + E[\text{var}(Z_t|Z_1, ..., Z_{t-1}, Y_t)]\\ =& \text{var}(Y_t|Z_1, ..., Z_{t-1}) + E[\tau^2|Z_1, ..., Z_{t-1}]\\ \equiv& P_{t|t-1} + \tau^2\end{aligned}$$

(*f*). With a similar argument as in (*a*) we also have that the combined vector $(\mathbf{Y},\mathbf{Z})$ is multivariate normal and therefore the conditional distribution in question is normal. The conditional means and variances are given directly from the definition. We further have that

$$\begin{aligned}\mathrm{cov}(Y_t,Z_t|Z_1,...,Z_{t-1}) =&\mathrm{cov}(E[Y_t|Z_1,...,Z_{t-1},Y_t],E[Z_t|Z_1,...,Z_{t-1},Y_t])+\\ &E[\mathrm{cov}(Y_t,Z_t|Z_1,...,Z_{t-1},Y_t)\\ =&\mathrm{cov}(Y_t,Y_t|Z_1,...,Z_{t-1})+0=P_{t|t-1}\end{aligned}$$

where the zero is obtained by that given Y_t Y_t itself is a constant and have zero variance. Based on this we have (using the rules for conditional normals)

$$\begin{aligned}\hat{y}_{t|t}\equiv&E(Y_t|Z_1,...,Z_{t-1},Z_t)\\ =&E[Y_t|Z_1,...,Z_{t-1}]+\frac{\mathrm{cov}[Y_t,Z_t|Z_1,...,Z_{t-1}]}{\mathrm{var}[Z_t|Z_1,...,Z_{t-1}]}(Z_t-E[Z_t|Z_1,...,Z_{t-1}])\\ =&\hat{y}_{t|t-1}+\frac{P_{t|t-1}}{S_t}(Z_t-\hat{z}_t)=\hat{y}_{t|t-1}+K_t(Z_t-\hat{z}_t)\\ P_{t|t}=&\mathrm{var}(Y_t|Z_1,...,Z_t)\\ =&\mathrm{var}(Y_t|Z_1,...,Z_{t-1})-\frac{[\mathrm{cov}[Y_t,Z_t|Z_1,...,Z_{t-1})]^2}{\mathrm{var}[Z_t|Z_1,...,Z_{t-1}]}\\ =&P_{t|t-1}-\frac{P^2_{t|t-1}}{S_t}=P_{t|t-1}[1-K_t]\end{aligned}$$

(*g*). We have that

$$\begin{aligned}\hat{y}_{t+1|t}=&E(Y_{t+1}|Z_1,...,Z_t)\\ =&E[E(Y_{t+1}|Z_1,...,Z_t,Y_t)]\\ =&E[\alpha Y_t|Z_1,...,Z_t)]\equiv\alpha\hat{y}_{t|t}\\ P_{t+1|t}=&\mathrm{var}(Y_{t+1}|Z_1,...,Z_t)\\ =&E[\mathrm{var}(Y_{t+1}|Z_1,...,Z_t,Y_t)+\mathrm{var}(E[Y_{t+1}|Z_1,...,Z_t,Y_t])\\ =&E[\sigma^2]+\mathrm{var}(\alpha Y_t|Z_1,...,Z_t)\\ =&\sigma^2+\alpha^2P_{t|t}\end{aligned}$$

(*h*). If Z_t is missing, then conditioning on $Z_1,...,Z_{t-1}$ is equivalent to conditioning on $Z_1,...,Z_t$ since Z_t contain no information. Thereby the result

(*i*). From the three equations we can calculate $\hat{z}_t$ and S_t recursively and thereby all the quantities involved for calculating the likelihood is available. Further, at each timestep we can multiply the previous value of the likelihood by $\phi(z_t;\hat{z}_t,S_t)$ making also the calculation of the likelihood recursively.

(*j*). The predictions seems to follow the observations but are closer to zero due to that the model has a prior prediction of zero on the process. When missing observations, the variance increases.

(*k*). The maximum value is about 0.895, not far from the true value. Note the smooth and unimodal behaviour of the likelihood function.

(*l*). The estimate of the observation error is far too small. This is reflected in that the predictions now follow the observations much closer. Further, the effect of missing observations is now much clearer in a larger increase in variance.

Solution 2 (to exercise 6) (*a*). Directly given by definition

(*b*). We have

$$\begin{aligned}
l(\mu,\sigma^2) =& -\tfrac{m}{2}\log(2\pi) - \tfrac{1}{2}\log|\sigma^2\mathbf{C}| - \frac{1}{2\sigma^2}(\mathbf{z}-\mu\mathbf{1})^T\mathbf{C}^{-1}(\mathbf{z}-\mu\mathbf{1})\\
=& -\tfrac{m}{2}\log(2\pi) - \tfrac{1}{2}\log|\sigma^2\mathbf{I}| - \tfrac{1}{2}\log|\mathbf{C}| - \frac{1}{2\sigma^2}(\mathbf{z}-\mu\mathbf{1})^T\mathbf{C}^{-1}(\mathbf{z}-\mu\mathbf{1})\\
=& -\tfrac{m}{2}\log(2\pi) - \tfrac{m}{2}\log(\sigma^2) - \tfrac{1}{2}\log|\mathbf{C}| - \frac{1}{2\sigma^2}(\mathbf{z}-\mu\mathbf{1})^T\mathbf{C}^{-1}(\mathbf{z}-\mu\mathbf{1})
\end{aligned}$$

giving

$$\frac{\partial}{\partial\mu}l(\mu,\sigma^2) = -\frac{1}{\sigma^2}(\mathbf{z}-\mu\mathbf{1})^T\mathbf{C}^{-1}\mathbf{1} = -\frac{1}{\sigma^2}\left[\mathbf{z}^T\mathbf{C}^{-1}\mathbf{1} - \mu\mathbf{1}\mathbf{C}^{-1}\mathbf{1}\right]$$

which, when putting to zero gives

$$\hat{\mu} = \frac{\mathbf{z}^T\mathbf{C}^{-1}\mathbf{1}}{\mathbf{1}\mathbf{C}^{-1}\mathbf{1}}$$

Further,

$$\frac{\partial}{\partial\sigma^2}l(\hat{\mu},\sigma^2) - \frac{m}{2\sigma^2} + \frac{1}{2\sigma^4}(\mathbf{z}-\hat{\mu}\mathbf{1})^T\mathbf{C}^{-1}(\mathbf{z}-\hat{\mu}\mathbf{1})$$

which, putting to zero gives

$$\hat{\sigma}^2 = \frac{1}{m}(\mathbf{z}-\hat{\mu}\mathbf{1})^T\mathbf{C}^{-1}(\mathbf{z}-\hat{\mu}\mathbf{1})$$

(*c*). For independent data, $\mathbf{C}=\mathbf{I}$, in which case

$$\begin{aligned}
\hat{\mu} =& \frac{\mathbf{z}\mathbf{1}}{\mathbf{1}\mathbf{1}} = \frac{1}{m}\sum_i z_i = \bar{z}\\
\hat{\sigma}^2 =& \frac{1}{m}(\mathbf{z}-\bar{z}\mathbf{1})^T(\mathbf{z}-\bar{z}\mathbf{1}) = \frac{1}{m}\sum_i (z_i-\bar{z})^2
\end{aligned}$$

(*d*). Since $\mathbf{Z} \sim \text{MVN}(\mu\mathbf{1}, \sigma^2\mathbf{C})$, we have $\widetilde{\mathbf{Z}} \sim \text{MVN}(\mu\mathbf{L1}, \sigma^2\mathbf{LCL}^T)$ and by using that $\mathbf{LCL}^T = \mathbf{LL}^{-1}(\mathbf{L}^T)^{-1}\mathbf{L}^T = \mathbf{I}$, we obtain the result. Then we have independent data with expectations dependening on one regression parameter μ.

(*e*). Using $\mathbf{X} = \mathbf{L1}$, we have ordinary estimates

$$\begin{aligned}\hat{\mu} =&(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\widetilde{\mathbf{Z}} = (\mathbf{1}^T\mathbf{L}^T\mathbf{L1})^{-1}\mathbf{1}^T\mathbf{L}^T\mathbf{LZ}\\ =&(\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1})^{-1}\mathbf{1}^T\mathbf{C}^{-1}\mathbf{Z}\\ \hat{\sigma}^2 =&\frac{1}{m}(\widetilde{\mathbf{Z}} - \hat{\mu}\mathbf{L1})^T(\widetilde{\mathbf{Z}} - \hat{\mu}\mathbf{L1}) = \frac{1}{m}(\mathbf{LZ} - \hat{\mu}\mathbf{L1})^T(\mathbf{LZ} - \hat{\mu}\mathbf{L1})\\ =&\frac{1}{m}(\mathbf{LZ} - \hat{\mu}\mathbf{1})^T\mathbf{L}^T\mathbf{L}(\mathbf{Z} - \hat{\mu}\mathbf{1}) = \frac{1}{m}(\mathbf{LZ} - \hat{\mu}\mathbf{1})^T\mathbf{C}^{-1}(\mathbf{Z} - \hat{\mu}\mathbf{1})\end{aligned}$$

(*f*). Weights 0.12718610.12718610.12657110.12657110.4924856. Reasonable that the close points are downweighted

Solution 3 (to exercise 20) (*a*). From exercise 12, we have that

$$Q_{ii}^{-1} = \text{var}[\delta_i|\delta_j, j \neq i] = \tau^2|\mathcal{N}(\mathbf{s}_i)|^{-1},$$

so $Q_{ii} = \tau^{-2}|\mathcal{N}(\mathbf{s}_i)|$.

Further, from exercise 12 (using that $\mu_i = 0$ in this case) we have that

$$E[\delta_i|\delta_j, j \neq i] = -Q_{ii}^{-1}\sum_{j\neq i} Q_{ij}\delta_j$$

which means that

$$Q_{ii}^{-1}Q_{ij} = \begin{cases}0 & \text{if } j \neq \mathcal{N}(\mathbf{s}_i)\\ -|\mathcal{N}(\mathbf{s}_i)|^{-1} & \text{if } j \in \mathcal{N}(\mathbf{s}_i)\end{cases} = -a_{ij}|\mathcal{N}(\mathbf{s}_i)|^{-1}$$

showing that

$$Q_{ij} = -a_{ij}\tau^{-2}$$

From $\mathbf{Q} = \mathbf{M}^{-1}(\mathbf{I} - \mathbf{H})$, we have that $Q_{ii} = M_{ii}^{-1}$ and $Q_{ij} = -M_{ii}^{-1}H_{ij}$ for $j \neq i$ which, by using the above results shows right.

(*b*). The ith element of $\mathbf{1}^T(\mathbf{I} - \mathbf{H})$ is equal to

$$1 - \sum_{j\neq i} H_{ij} = 1 - \sum_{j\neq i} a_{ij}|\mathcal{N}(\mathbf{s}_i)|^{-1} = 1 - \sum_{j\in\mathcal{N}(\mathbf{s}_i)} |\mathcal{N}(\mathbf{s}_i)|^{-1} = 1 - 1 = 0$$

This means that zero is an eigenvalue and the matrix is singular. Then also $\mathbf{Q}$ becomes singular making it impossible to invert.

(c). We have that if $\boldsymbol{\delta}$ is Gaussian, then also $\boldsymbol{\xi}$ is Gaussian. Further

$$
\begin{aligned}
E[\xi_i|\xi_j, j\neq i] =& E[\delta_i + c|\delta_j + c, j\neq i]\\
=& c + E[\delta_i|\delta_j, j\neq i]\\
=& c + |\mathcal{N}(\mathbf{s}_i)|^{-1}\sum_{j\in\mathcal{N}(\mathbf{s}_i)}\delta_j\\
=& |\mathcal{N}(\mathbf{s}_i)|^{-1}\sum_{j\in\mathcal{N}(\mathbf{s}_i)}[\delta_j + c]\\
=& |\mathcal{N}(\mathbf{s}_i)|^{-1}\sum_{j\in\mathcal{N}(\mathbf{s}_i)}[\xi_j]\\
\mathrm{var}[\xi_i|\xi_j, j\neq i] =& \mathrm{var}[\delta_i + c|\delta_j + c, j\neq i]\\
=& \mathrm{var}[\delta_i|\delta_j, j\neq i]\\
=& \tau^2|\mathcal{N}(\mathbf{s}_i)|^{-1}
\end{aligned}
$$

showing that $\boldsymbol{\xi}$ has the same conditional distributions as $\boldsymbol{\delta}$. Therefore it is not possible to distinguish between $\boldsymbol{\delta}$ and $\boldsymbol{\xi}$ through the conditional distributions.

(d). We have

$$
\begin{aligned}
p(\boldsymbol{\delta}|\mathbf{Z}) \propto& p(\boldsymbol{\delta})p(\mathbf{Z}|\boldsymbol{\delta})\\
&\propto \exp(-0.5\boldsymbol{\delta}^T\mathbf{Q}\boldsymbol{\delta})\prod_{i=1}^m \exp(-0.5\sigma^{-2}(Z_i-\delta_i)^2)\\
&\propto \exp(-0.5\boldsymbol{\delta}^T\mathbf{Q}\boldsymbol{\delta})\exp(-0.5\sigma^{-2}(\mathbf{Z}-\boldsymbol{\delta})^T(\mathbf{Z}-\boldsymbol{\delta})\\
&\propto \exp(-0.5[\boldsymbol{\delta}^T[\mathbf{Q}+\sigma^{-2}\mathbf{I}]\boldsymbol{\delta} - 2\sigma^{-2}\mathbf{Z}^T\boldsymbol{\delta}])\\
&\propto \exp(-0.5[(\boldsymbol{\delta}^T - \sigma^{-2}[\mathbf{Q}+\sigma^{-2}\mathbf{I}]^{-1}\mathbf{Z})^T[\mathbf{Q}+\sigma^{-2}\mathbf{I}](\boldsymbol{\delta} - \sigma^{-2}[\mathbf{Q}+\sigma^{-2}\mathbf{I}]^{-1}\mathbf{Z})])
\end{aligned}
$$

showing that $[\boldsymbol{\delta}|\mathbf{Z}]$ has a precision matrix equal to $\mathbf{Q}+\sigma^{-2}\mathbf{I}$. Further,

$$
\begin{aligned}
\mathbf{Q}+\sigma^{-2}\mathbf{I} =& \tilde{\tau}^2(\mathbf{I}-\mathbf{H}) + \sigma^{-2}\mathbf{I}\\
=& (\tilde{\tau}^2+\sigma^{-2})\mathbf{I} - \tilde{\tau}^2\mathbf{H}\\
=& (\tilde{\tau}^2+\sigma^{-2})[\mathbf{I} - \frac{\tilde{\tau}^2}{\tilde{\tau}^2+\sigma^{-2}}\mathbf{H}]
\end{aligned}
$$

See script for numerical evaluation of eigenvalues

(e). We have

$$
\begin{aligned}
\mathrm{var}[\delta_i|\delta_j, j\neq i] =& Q_{ii}^{-1} = [\tau^{-2}\Delta_{ii}^{-1}]^{-1} = \tau^2|\mathcal{N}(\mathbf{s}_i)|^{-1}\\
E[\delta_i\delta_j, j\neq i] =& -Q_{ii}^{-1}\sum_{j\neq i}Q_{ij}\delta_j = [\tau^{-2}\Delta_{ii}^{-1}]^{-1}\sum_{j\in\mathcal{N}(\mathbf{s}_i)}[\tau^{-2}\Delta_{ii}^{-1}]\phi|\mathcal{N}(\mathbf{s}_i)|^{-1}\delta_j\\
=& \frac{\phi}{|\mathcal{N}(\mathbf{s}_i)|}\sum_{j\in\mathcal{N}(\mathbf{s}_i)}\delta_j
\end{aligned}
$$

(f). We have

$$\text{var}[\widetilde{\boldsymbol{\delta}}] = \boldsymbol{\Delta}^{1/2}\text{var}[\boldsymbol{\delta}]\boldsymbol{\Delta}^{1/2} = \boldsymbol{\Delta}^{1/2}[\tau^{-2}\boldsymbol{\Delta}^{-1}[\mathbf{I} - \phi\mathbf{H}]]^{-1}\boldsymbol{\Delta}^{1/2}$$

The precision matrix for $\widetilde{\boldsymbol{\delta}}$ is then

$$\begin{aligned}\widetilde{\mathbf{Q}} =& \boldsymbol{\Delta}^{-1/2}[\tau^{-2}\boldsymbol{\Delta}^{-1}[\mathbf{I} - \phi\mathbf{H}]]\boldsymbol{\Delta}^{-1/2}\\ =& \tau^{-2}[\mathbf{I} - \phi\boldsymbol{\Delta}^{-1/2}\mathbf{H}\boldsymbol{\Delta}^{1/2}] = \tau^{-2}[\mathbf{I} - \phi\widetilde{\mathbf{H}}]\end{aligned}$$

(g). Given directly from $\boldsymbol{\delta} = \boldsymbol{\Delta}^{-1/2}\widetilde{\boldsymbol{\delta}}$.

(h). The ***generic1*** model is

$$\begin{aligned}\mathbf{Q} =& \xi(\mathbf{I} - \tfrac{\beta}{\lambda_{max}}\widetilde{\mathbf{H}})\\ , \phi \in& (0,1), \lambda_{\max} \text{ maximum eigenvalue of } \widetilde{\mathbf{H}}.\end{aligned}$$

Here $\lambda_{\max} = 0.0008132974$ while $\hat{\beta} = 4.942e - 01$ which gives $\hat{\phi} = \hat{\beta} * \lambda_{\max} = 0.0004019316$.

Solution 4 (to exercise 21) (a). We have

$$\begin{aligned}p(\mu|\tau, \mathbf{y}) \propto& p(\mu)p(\mathbf{y}|\mu, \tau)\\ &\propto \exp(-0.5\tau_0(\mu - \mu_0)^2)\tau^{n/2}\prod_{i=1}^{n}\exp(-0.5\tau(y_i - \mu)^2)\\ &\propto \exp(-0.5([\tau_0 + n\tau]\mu^2 - 2[\tau_0\mu_0 + n\tau\bar{y}]\mu)\\ &\propto \exp(-0.5[\tau_0 + n\tau](\mu^2 - 2\frac{[\tau_0\mu_0 + n\tau\bar{y}]}{[\tau_0 + n\tau]}\mu)\\ &\propto \exp(-0.5[\tau_0 + n\tau](\mu - \frac{[\tau_0\mu_0 + n\tau\bar{y}]}{[\tau_0 + n\tau]})^2\\ &\propto N\left(\frac{[\tau_0\mu_0 + n\tau\bar{y}]}{[\tau_0 + n\tau]}, [\tau_0 + n\tau]^{-1}\right)\end{aligned}$$

Similarly

$$\begin{aligned}p(\tau|\mu, \mathbf{y}) \propto& p(\tau)p(\mathbf{y}|\mu, \tau)\\ &\propto \tau^{a-1}\exp(-b\tau)\tau^{n/2}\prod_{i=1}^{n}\exp(-0.5\tau(y_i - \mu)^2)\\ &\propto \tau^{a+n/2-1}\exp(-\tau[b + 0.5\sum_{i=1}^{n}(y_i - \mu)^2])\\ &\propto \tau^{a+n/2-1}\exp(-\tau[b + 0.5\sum_{i=1}^{n}(y_i - \bar{y})^2 + n(\bar{y} - \mu)^2])\\ &\propto \text{Gamma}(a + n/2, b + 0.5\sum_{i=1}^{n}(y_i - \mu)^2)\end{aligned}$$

Note that $\sum_{i=1}^n (y_i-\mu)^2 = \sum_{i=1}^n (y_i-\bar{y})^2 + n(\bar{y}-\mu)^2$. From this we see that the data only enter the conditional distributions through $\bar{y}$ and $\sum_{i=1}^n (y_i-\bar{y})^2$.

(*b*). See script

(*c*). See script

Solution 5 (to 22) (*a*). We have that

$$\begin{aligned}\Pr(Y_s = y_s|\mathbf{Y}_u = y_u, u\neq s) \propto& \exp(\exp\{-\beta\sum_{v\sim u} y_u y_v\})\\ \propto& \exp(\exp\{-\beta\sum_{v\sim s} y_u y_s\})\end{aligned}$$

showing that $\mathcal{N}(s)$ is the set of the four nearest neighbors (with some corrections on the borders).

(*b*). We do not observe y_u "directly" but only a smoothed (blurred) version of y_u where the bluring is an average of the variable it self and the nearest neighbors.

(*c*). We now have

$$\begin{aligned}\Pr(\mathbf{y}|\mathbf{z}) \propto& \Pr(\mathbf{y})\Pr(\mathbf{z}|\mathbf{y})\\ \propto& \exp\{-\beta\sum_{v\sim u} y_u y_v\}\prod_u f_e(z_u - 5^{-1}\sum_{t\in\nu(u)} y_t)\\ \propto& \exp\{-\beta\sum_{v\sim u} y_u y_v + \sum_u \log f_e(z_u - 5^{-1}\sum_{t\in\nu(u)} y_t)\}\end{aligned}$$

This shows that now $\nu(u)$ is a clique. This again means that all members of $\nu(u)$ are neighbors. This then gives the neighborhood set ($u=(i,j)$) $\mathcal{N}(u) = \{(i-2,j),(i-1,j-1),(i-1,j),(i-1,j+1),(i,j-2),(i,j-1),(i,j+1),(i,j+2),(i+1,j-1),(i+1,j),(i+1,j+1),(i,j+2)\}$

Cliques:

Largest clique is $\{(i,j),(i,j+1),(i+1,j),(i-1,j),(i,j-1)\}$.

Largest clique not a subset of the above is $\{(i,j),(i,j+1),(i+1,j),(i+1,j+1)\}$

Any subset of these "maximal" cliques are also cliques.

- Single points: $\{(i,j)\}$
- Pairs $\{(i,j),(i,j+1)\}$, $\{(i,j),(i+1,j)\}$, $\{(i,j),(i,j+2)\}$, $\{(i,j),(i+2,j)\}$, $\{(i,j),(i+1,j+1)\}$,

- Triplets $\{(i,j),(i,j+1),(i+1,j)\}$, $\{(i,j),(i,j+1),(i+1,j+1)\}$, $\{(i,j),(i,j+1),(i,j+2)\}$, $\{(i,j),(i,j+1),(i-1,j)\}$, $\{(i,j),(i+1,j),(i+2,j)\}$, $\{(i,j),(i+1,j),(i+1,j-1)\}$, $\{(i,j),(i+1,j),(i-1,j)\}$
- Four $\{(i,j),(i,j+1),(i+1,j),(i+1,j+1)\}$, $\{(i,j),(i,j+1),(i+1,j),(i-1,j)\}$, $\{(i,j),(i,j-1),(i+1,j),(i-1,j)\}$, $\{(i,j),(i+1,j),(i+2,j),(i+1,j+1)\}$, $\{(i,j),(i+1,j),(i+2,j),(i+1,j-1)\}$
- Five $\{(i,j),(i,j+1),(i+1,j),(i-1,j),(i,j-1)\}$

Solution 6 (to 28) (a).

(b).

(c). This exercise turned out to be a bit difficult. It is not required that you understand this, but a solution (thanks to Dag Sverre Seljebotn) is included for those of interest.

Assume $\mathbf{A}$ and $\mathbf{B}$ are symmetric matrices of dimensions $n\times n$ and $m\times m$ respectivly, and both are positive definite. We want to show that $\mathbf{A}\otimes\mathbf{B}$ also is positive definite. Since $\mathbf{B}$ is positive definite, there exist a (positive definite) matrix $\mathbf{L}$ such that $\mathbf{B}=\mathbf{L}\mathbf{L}^T$. Define further $\mathbf{D}=\mathbf{I}_m\otimes\mathbf{L}$. Then

$$\begin{aligned}\mathbf{D}[\mathbf{A}\otimes\mathbf{I}_n]\mathbf{D}^T=&[\mathbf{I}_n\otimes\mathbf{L}][\mathbf{A}\otimes\mathbf{I}_n][\mathbf{I}_n\otimes\mathbf{L}^T]\\=&\mathbf{I}_n\mathbf{A}\mathbf{I}_n\otimes(\mathbf{L}\mathbf{L}^T)=\mathbf{A}\otimes\mathbf{B}\end{aligned}$$

Then

$$\begin{aligned}\mathbf{x}^T[\mathbf{A}\otimes\mathbf{B}]\mathbf{x}=&\mathbf{x}^T\mathbf{D}[\mathbf{A}\otimes\mathbf{I}_n]\mathbf{D}^T\mathbf{x}\\=&\tilde{\mathbf{x}}^T[\mathbf{A}\otimes\mathbf{I}_n]\tilde{\mathbf{x}}\end{aligned}$$

where $\tilde{\mathbf{x}}=\mathbf{D}^T\mathbf{x}$. It is therefore sufficient to show that $\mathbf{A}\otimes\mathbf{I}_n$ is positive definite. Now define $\mathbf{P}=\frac{1}{\sqrt{m}}\mathbf{1}\otimes\mathbf{A}$ where $\mathbf{1}$ now is an $m\times m$ matrix only consisting of ones. Then (check this yourself)

$$\mathbf{A}\otimes\mathbf{I}_n=\mathbf{P}[\mathbf{I}_m\otimes\mathbf{A}]\mathbf{P}^T$$

Then

$$\begin{aligned}\tilde{\mathbf{x}}^T[\mathbf{A}\otimes\mathbf{I}_n]\tilde{\mathbf{x}}=&\tilde{\mathbf{x}}^T\mathbf{P}[\mathbf{I}_m\otimes\mathbf{A}]\mathbf{P}^T\tilde{\mathbf{x}}\\=&\hat{\mathbf{x}}^T[\mathbf{I}_m\otimes\mathbf{A}]\hat{\mathbf{x}}\end{aligned}$$

where $\hat{\mathbf{x}}=\mathbf{P}^T\tilde{\mathbf{x}}$. Now writing $\hat{\mathbf{x}}=(\hat{\mathbf{x}}_1,...,\hat{\mathbf{x}}_m)'$, we get

$$\hat{\mathbf{x}}^T[\mathbf{I}_m\otimes\mathbf{A}]\hat{\mathbf{x}}=\sum_{k=1}^m\hat{\mathbf{x}}_k\mathbf{A}\hat{\mathbf{x}}_k\geq 0$$

and the result follows.