Lecture 23

Spatio-temporal Models

Colin Rundel

04/17/2017

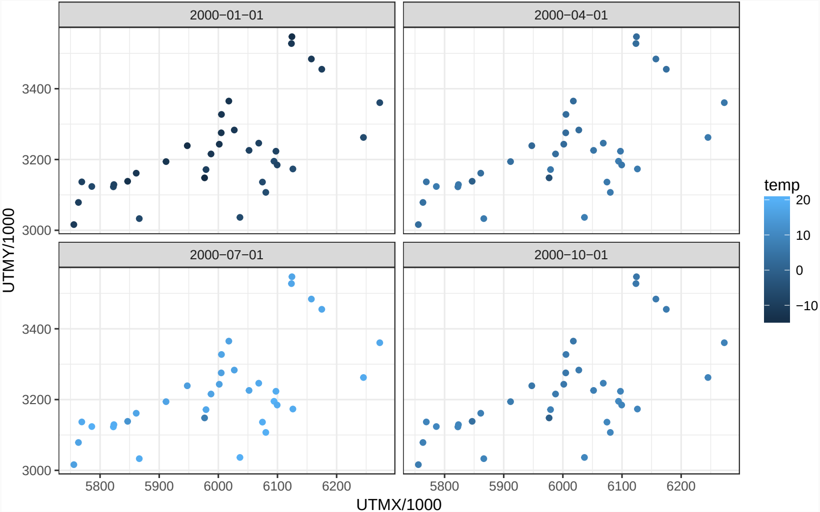
Spatial Models with AR time dependence

Example - Weather station data

Based on Andrew Finley and Sudipto Banerjee's notes from National Ecological Observatory Network (NEON) Applied Bayesian Regression Workshop, March 7 - 8, 2013 Module 6

`NETemp.dat` - Monthly temperature data (Celsius) recorded across the Northeastern US starting in January 2000.

```
library(spBayes)
data("NETemp.dat")
ne_temp = NETemp.dat %>%
  filter(UTMX > 5.5e6, UTMY > 3e6) %>%
  select(1:27) %>%
  tbl_df()
ne_temp
## # A tibble: 34 × 27
##     elev    UTMX    UTMY        y.1       y.2        y.3      y.4       y.5
##    <int>   <dbl>   <dbl>      <dbl>     <dbl>      <dbl>    <dbl>     <dbl>
## 1    102 6094162 3195181  -6.388889 -3.611111  3.7222222 6.777778 12.555556
## 2      1 6245390 3262354  -6.277778 -4.111111  2.6111111 6.555556 11.388889
## 3    157 6157302 3484043 -11.111111 -9.444444 -0.3888889 3.944444  9.888889
## 4    176 6123610 3527665 -11.611111 -9.722222 -1.1666667 2.888889  9.666667
## 5    400 6004871 3275456 -12.611111 -9.055556 -1.6111111 2.555556  8.555556
## 6    133 6051946 3225830  -9.111111 -6.388889  1.2222222 4.944444 10.888889
## 7     56 6099462 3184587  -7.944444 -6.055556  2.0555556 5.555556 11.111111
## 8     59 6074601 3136288  -6.555556 -3.500000  3.1666667 6.166667 11.500000
## 9    160 6174891 3455064  -9.944444 -8.944444 -0.2777778 3.555556  9.611111
## 10   360 6005282 3327413 -12.277778 -9.444444 -1.5000000 2.944444  9.000000
## # ... with 24 more rows, and 19 more variables: y.6 <dbl>, y.7 <dbl>,
## #   y.8 <dbl>, y.9 <dbl>, y.10 <dbl>, y.11 <dbl>, y.12 <dbl>, y.13 <dbl>,
```


2000−01−01
2000−04−01
2000−07−01
2000−10−01
3400
3200
3000
5800
5900
6000
6100
6200
UTMY/1000
UTMX/1000
temp
20
10
0
−10

$$y_t = \underset{1\times p}{F'_t} \underset{p\times 1}{\boldsymbol{\theta}_t} + v_t \qquad \text{observation equation}$$

$$\underset{p\times 1}{\boldsymbol{\theta}_t} = \underset{p\times p}{G_t} \underset{p\times 1}{\boldsymbol{\theta}}_{t-1} + \underset{p\times 1}{\boldsymbol{\omega}_t} \qquad \text{evolution equation}$$

$$\boldsymbol{v}_t \sim \mathcal{N}(0, \boldsymbol{V}_t)$$
$$\boldsymbol{\omega}_t \sim \mathcal{N}(0, \boldsymbol{W}_t)$$

ARMA / ARIMA are a special case of a dynamic linear model, for example an $AR(p)$ can be written as

$$F_t' = (1, 0, \ldots, 0)$$

$$G_t = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

$$\omega_t = (\omega_1, 0, \ldots 0), \qquad \omega_1 \sim \mathcal{N}(0,\ \sigma^2)$$

DLM vs ARMA

ARMA / ARIMA are a special case of a dynamic linear model, for example an $AR(p)$ can be written as

$$F_t' = (1, 0, \ldots, 0)$$

$$G_t = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

$$\omega_t = (\omega_1, 0, \ldots 0), \qquad \omega_1 \sim \mathcal{N}(0,\ \sigma^2)$$

$$\begin{aligned} y_t &= \theta_t + v_t, & v_t &\sim \mathcal{N}(0,\ \sigma_v^2) \\ \theta_t &= \sum_{i=1}^{p} \phi_i\, \theta_{t-i} + \omega_1, & \omega_1 &\sim \mathcal{N}(0,\ \sigma_\omega^2) \end{aligned}$$

The observed temperature at time t and location s is given by $y_t(s)$ where,

$$
\begin{aligned}
y_t(s) &= x_t(s)\boldsymbol{\beta}_t + u_t(s) + \epsilon_t(s) \\
\epsilon_t(s) &\overset{ind.}{\sim} \mathcal{N}(0, \tau_t^2)
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{\beta}_t &= \boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t \\
\boldsymbol{\eta}_t &\overset{i.i.d.}{\sim} \mathcal{N}(0, \boldsymbol{\Sigma}_\eta)
\end{aligned}
$$

$$
\begin{aligned}
u_t(s) &= u_{t-1}(s) + w_t(s) \\
w_t(s) &\overset{ind.}{\sim} \mathcal{N}\left(\mathbf{0}, \Sigma_t(\phi_t, \sigma_t^2)\right)
\end{aligned}
$$

The observed temperature at time t and location s is given by $y_t(s)$ where,

$$
\begin{aligned}
y_t(s) &= x_t(s)\boldsymbol{\beta}_t + u_t(s) + \epsilon_t(s) \\
\epsilon_t(s) &\overset{ind.}{\sim} \mathcal{N}(0, \tau_t^2)
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{\beta}_t &= \boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t \\
\boldsymbol{\eta}_t &\overset{i.i.d.}{\sim} \mathcal{N}(0, \boldsymbol{\Sigma}_\eta)
\end{aligned}
$$

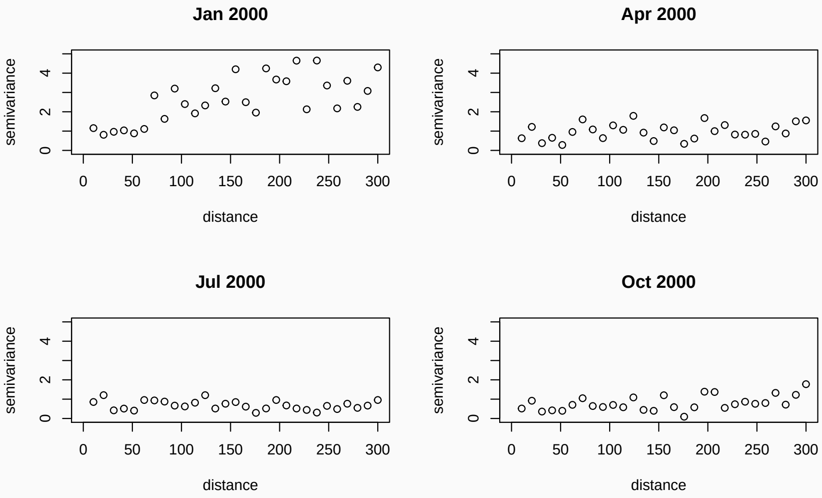
$$
\begin{aligned}
u_t(s) &= u_{t-1}(s) + w_t(s) \\
w_t(s) &\overset{ind.}{\sim} \mathcal{N}\left(\mathbf{0}, \Sigma_t(\phi_t, \sigma_t^2)\right)
\end{aligned}
$$

Additional assumptions for $t = 0$,

$$\boldsymbol{\beta}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$$

$$u_0(s) = 0$$

Variograms by time



Data and Model Parameters

**Data*:

```
coords = ne_temp %>% select(UTMX, UTMY) %>% as.matrix() / 1000
y_t = ne_temp %>% select(starts_with("y.")) %>% as.matrix()

max_d = coords %>% dist() %>% max()
n_t = ncol(y_t)
n_s = nrow(y_t)
```

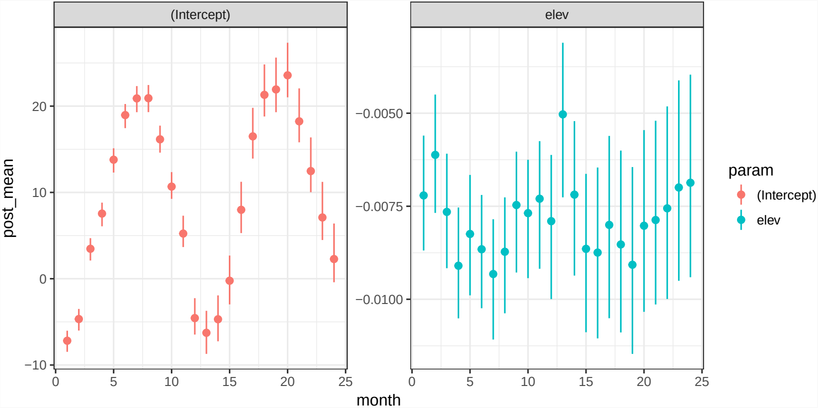
**Parameters*:

```
n_beta = 2
starting = list(
  beta = rep(0, n_t * n_beta), phi = rep(3/(max_d/2), n_t),
  sigma.sq = rep(1, n_t), tau.sq = rep(1, n_t),
  sigma.eta = diag(0.01, n_beta)
)

tuning = list(phi = rep(1, n_t))

priors = list(
  beta.0.Norm = list(rep(0, n_beta), diag(1000, n_beta)),
  phi.Unif = list(rep(3/(0.9 * max_d), n_t), rep(3/(0.05 * max_d), n_t)),
  sigma.sq.IG = list(rep(2, n_t), rep(2, n_t)),
  tau.sq.IG = list(rep(2, n_t), rep(2, n_t)),
  sigma.eta.IW = list(2, diag(0.001, n_beta))
)
```

Fitting with spDynLM from spBayes

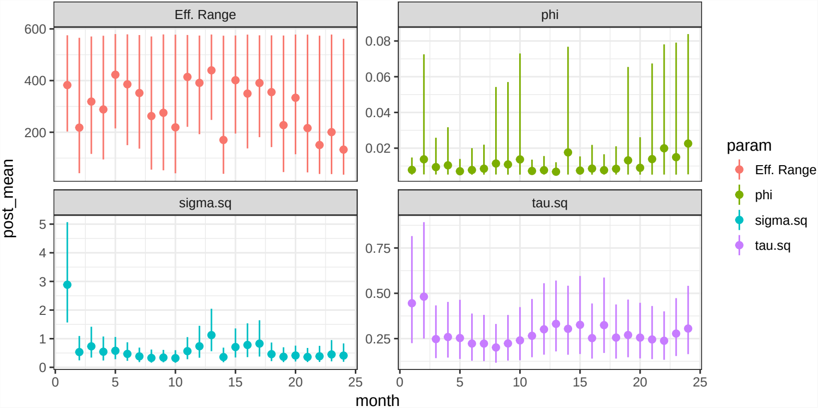
```
n_samples = 10000
models = lapply(paste0("y.",1:24, "~elev"), as.formula)

m = spDynLM(
  models, data = ne_temp, coords = coords, get.fitted = TRUE,
  starting = starting, tuning = tuning, priors = priors,
  cov.model = "exponential", n.samples = n_samples, n.report = 1000)

save(m, ne_temp, models, coords, starting, tuning, priors, n_samples,
     file="dynlm.Rdata")

## ----------------------------------------
##     General model description
## ----------------------------------------
## Model fit with 34 observations in 24 time steps.
##
## Number of missing observations 0.
##
## Number of covariates 2 (including intercept if specified).
##
## Using the exponential spatial correlation model.
##
## Number of MCMC samples 10000.
##
## ...
```

Posterior Inference - βs

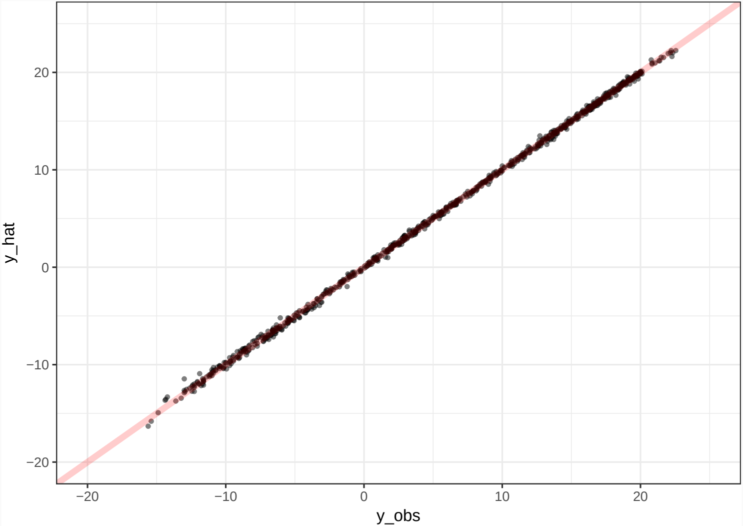


Lapse Rate

Posterior Inference - θ



Posterior Inference - Observed vs. Predicted

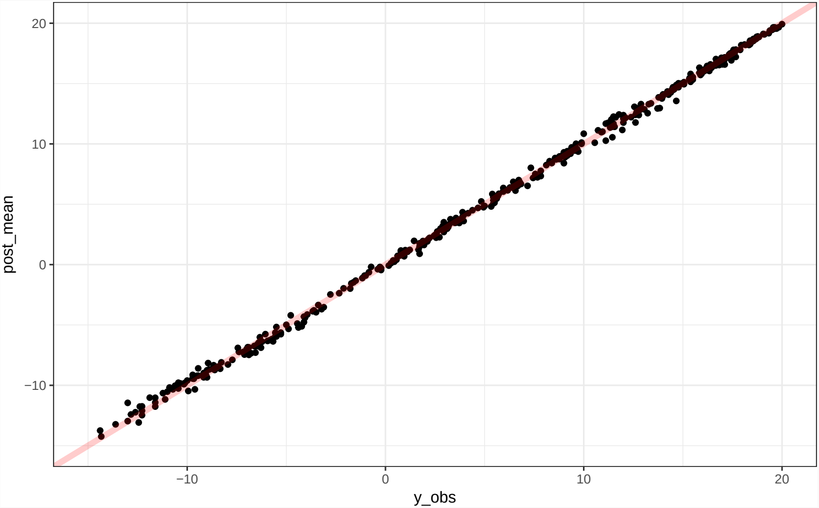


Prediction

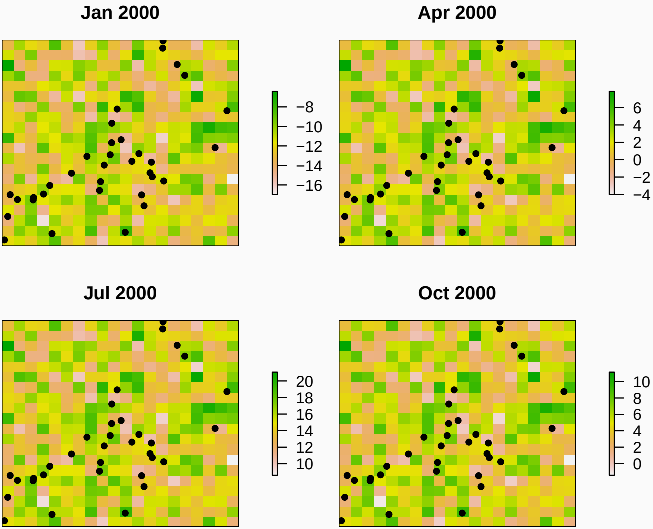
`spPredict` does not support `spDynLM` objects.

```
r = raster(xmn=575e4, xmx=630e4, ymn=300e4, ymx=355e4, nrow=20, ncol=20)

pred = xyFromCell(r, 1:length(r)) %>%
  cbind(elev=0, ., matrix(NA, nrow=length(r), ncol=24)) %>%
  as.data.frame() %>%
  setNames(names(ne_temp)) %>%
  rbind(ne_temp, .) %>%
  select(1:15) %>%
  select(-elev)
```

```
models_pred = lapply(paste0("y.",1:n_t, "~1"), as.formula)

n_samples = 5000
m_pred = spDynLM(
  models_pred, data = pred, coords = coords_pred, get.fitted = TRUE,
  starting = starting, tuning = tuning, priors = priors,
  cov.model = "exponential", n.samples = n_samples, n.report = 1000)

save(m_pred, pred, models_pred, coords_pred, y_t_pred, n_samples,
     file="dynlm_pred.Rdata")
```



post_mean
y_obs
20
10
0
−10
−10
0
10
20


Jan 2000
−8
−10
−12
−14
−16
Apr 2000
6
4
2
0
−2
−4
Jul 2000
20
18
16
14
12
10
Oct 2000
10
8
6
4
2
0

Spatio-temporal models for continuous time

Additive Models

In general, spatiotemporal models will have a form like the following,

$$
\begin{aligned}
y(\mathbf{s},t) &= \underset{\text{mean structure}}{\mu(\mathbf{s},t)} + \underset{\text{error structure}}{e(\mathbf{s},t)} \\
&= \underset{\text{Regression}}{\mathbf{x}(\mathbf{s},t)\,\boldsymbol{\beta}(\mathbf{s},t)} + \underset{\text{Spatiotemporal RE}}{w(\mathbf{s},t)} + \underset{\text{Error}}{\epsilon(\mathbf{s},t)}
\end{aligned}
$$

Additive Models

In general, spatiotemporal models will have a form like the following,

$$
\begin{aligned}
y(\mathbf{s},t) &= \underset{\text{mean structure}}{\mu(\mathbf{s},t)} + \underset{\text{error structure}}{e(\mathbf{s},t)} \\
&= \underset{\text{Regression}}{\mathbf{x}(\mathbf{s},t)\,\boldsymbol{\beta}(\mathbf{s},t)} + \underset{\text{Spatiotemporal RE}}{w(\mathbf{s},t)} + \underset{\text{Error}}{\epsilon(\mathbf{s},t)}
\end{aligned}
$$

The simplest possible spatiotemporal model is one were assume there is no dependence between observations in space and time,

$$
w(\mathbf{s},t) = \alpha(t) + \omega(\mathbf{s})
$$

these are straight forward to fit and interpret but are quite limiting (no shared information between space and time).

Spatiotemporal Covariance

Lets assume that we want to define our spatiotemporal random effect to be a single stationary Gaussian Process (in 3 dimensions*),

$$w(s,t) \sim \mathcal{N}\left(0, \boldsymbol{\Sigma}(s,t)\right)$$

where our covariance function depends on both $\|s - s'\|$ and $|t - t'|$,

$$\text{cov}(w(s,t), w(s',t')) = c(\|s - s'\|, |t - t'|)$$

- Note that the resulting covariance matrix Σ will be of size $n_s \cdot n_t \times n_s \cdot n_t$.
  - Even for modest problems this gets very large (past the point of direct computability).
  - If $n_t = 52$ and $n_s = 100$ we have to work with a 5200×5200 covariance matrix

Separable Models

One solution is to use a seperable form, where the covariance is the product of a valid 2d spatial and a valid 1d temporal covariance / correlation function,

$$\text{cov}(w(s,t), w(s',t')) = \sigma^2 \, \rho_1(\|s - s'\|; \boldsymbol{\theta}) \, \rho_2(|t - t'|; \boldsymbol{\phi})$$

One solution is to use a seperable form, where the covariance is the product of a valid 2d spatial and a valid 1d temporal covariance / correlation function,

$$\text{cov}(\boldsymbol{w}(\boldsymbol{s}, \boldsymbol{t}), \boldsymbol{w}(\boldsymbol{s}', \boldsymbol{t}')) = \sigma^2 \, \rho_1(\|\boldsymbol{s} - \boldsymbol{s}'\|; \boldsymbol{\theta}) \, \rho_2(|\boldsymbol{t} - \boldsymbol{t}'|; \boldsymbol{\phi})$$

If we define our observations as follows (stacking time locations within spatial locations)

$$\boldsymbol{w}_{\boldsymbol{s}}^{t} = \Big(w(s_1, t_1), \cdots, w(s_1, t_{n_t}), w(s_2, t_1), \cdots, w(s_2, t_{n_t}), \cdots, \cdots, w(s_{n_s}, t_1), \cdots, w(s_{n_s}, t_{n_t})\Big)$$

then the covariance can be written as

$$\boldsymbol{\Sigma}_w(\sigma^2, \theta, \phi) = \sigma^2 \, H_s(\theta) \otimes H_t(\phi)$$

where $H_s(\theta)$ and $H_t(\theta)$ are $n_s \times n_s$ and $n_t \times n_t$ sized correlation matrices respectively and their elements are defined by

$$\{H_s(\theta)\}_{ij} = \rho_1(\|\boldsymbol{s}_i - \boldsymbol{s}_j\|; \theta)$$
$$\{H_t(\phi)\}_{ij} = \rho_1(|t_i - t_j|; \phi)$$

Kronecker Product

Definition:

$$\underset{[m\times n]}{A} \otimes \underset{[p\times q]}{B} = \underset{[m\cdot p\times n\cdot q]}{\begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}}$$

Kronecker Product

Definition:

$$\underset{[m\times n]}{A} \otimes \underset{[p\times q]}{B} = \underset{[m\cdot p\times n\cdot q]}{\begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}}$$

Properties:

$$\begin{aligned} A \otimes B &\neq B \otimes A \qquad \text{(usually)} \\ (A \otimes B)^t &= A^t \otimes B^t \\ \det(A \otimes B) &= \det(B \otimes A) \\ &= \det(A)^{\operatorname{rank}(B)} \det(B)^{\operatorname{rank}(A)} \\ (A \otimes B)^{-1} &= A^{-1}B^{-1} \end{aligned}$$

If we have a spatiotemporal random effect with a separable form,

$$w(s,t) \sim \mathcal{N}(0,\ \boldsymbol{\Sigma}_w)$$

$$\boldsymbol{\Sigma}_w = \sigma^2\, H_s \otimes H_t$$

then the likelihood for $\boldsymbol{w}$ is given by

$$
\begin{aligned}
& -\frac{n}{2}\log 2\pi - \frac{1}{2}\log|\Sigma_w| - \frac{1}{2}\boldsymbol{w}^t \boldsymbol{\Sigma}_{\boldsymbol{w}}^{-1}\boldsymbol{w} \\
= & -\frac{n}{2}\log 2\pi - \frac{1}{2}\log\left[(\sigma^2)^{n_t \cdot n_s}|H_s|^{n_t}|H_t|^{n_s}\right] - \frac{1}{2}\boldsymbol{w}^t\frac{1}{\sigma^2}(\boldsymbol{H}_s^{-1} \otimes \boldsymbol{H}_t^{-1})\boldsymbol{w}
\end{aligned}
$$

Non-seperable Models

- Additive and separable models are still somewhat limiting
- Cannot treat spatiotemporal covariances as 3d observations
- Possible alternatives:
  - Specialized spatiotemporal covariance functions, i.e.

    $$c(\mathbf{s}-\mathbf{s}', t-t') = \sigma^2(|t-t'|+1)^{-1} \exp\left(-\|\mathbf{s}-\mathbf{s}'\|(|t-t'|+1)^{-\beta/2}\right)$$

  - Mixtures, i.e. $w(\mathbf{s}, t) = w_1(\mathbf{s}, t) + w_2(\mathbf{s}, t)$, where $w_1(\mathbf{s}, t)$ and $w_2(\mathbf{s}, t)$ have seperable forms.