A structured random effects models introdution, the GMRFs

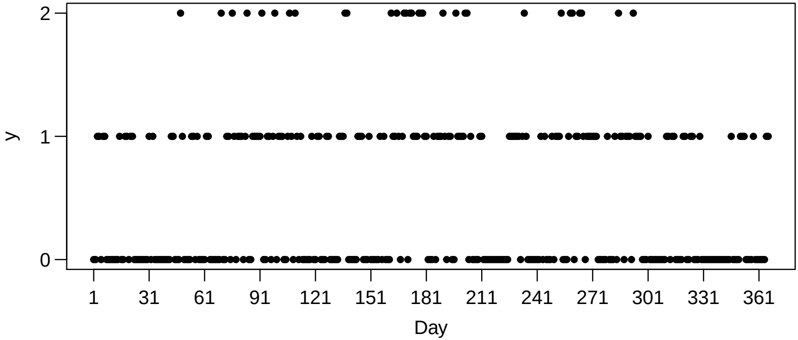
INLA team

Last update, November 2019

Motivating examples

Tokyo example

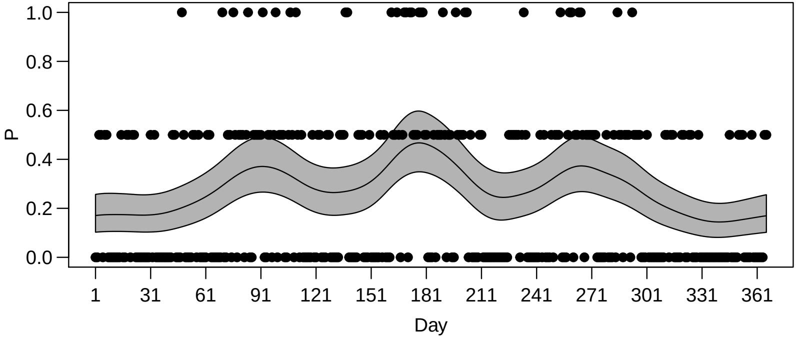
Rainfall over 1 mm in the Tokyo area for each calendar day during two years (1983-84) are registered.



Problem: model the probability of rain each day of the year

Tokyo example (fit)

Rainfall over 1 mm in the Tokyo area for each calendar day during two years (1983-84) are registered.



The probability of rain each day of the year

Scotland example

1. Conditional Autoregressive (CAR) models for disease mapping: Lip cancer in Scotland [top]

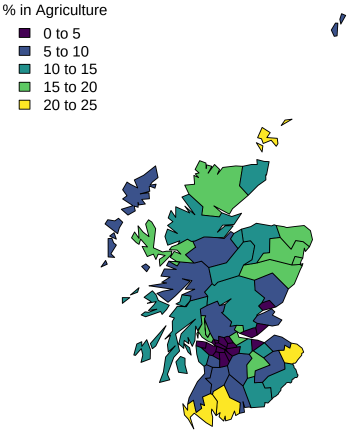
The rates of lip cancer in 56 counties in Scotland have been analysed by Clayton and Kaldor (1987) and Breslow and Clayton (1993). The form of the data includes the observed and expected cases (expected numbers based on the population and its age and sex distribution in the county), a covariate measuring the percentage of the population engaged in agriculture, fishing, or forestry, and the "position" of each county expressed as a list of adjacent counties.

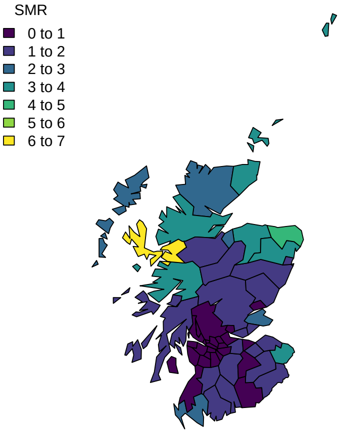
County	Observed cases O_i	Expected cases E_i	Percentage in agric. x_i	SMR	Adjacent counties
1	9	1.4	16	652.2	5,9,11,19
2	39	8.7	16	450.3	7,10
...	...	...	...	...	...
56	0	1.8	10	0.0	18,24,30,33,45,55

We note that the extreme SMRs (Standardised Mortality Ratios) are based on very few cases.

Figure 1: Scootland example from WinBUGS.

Scotland maps





Scotland data: GLM

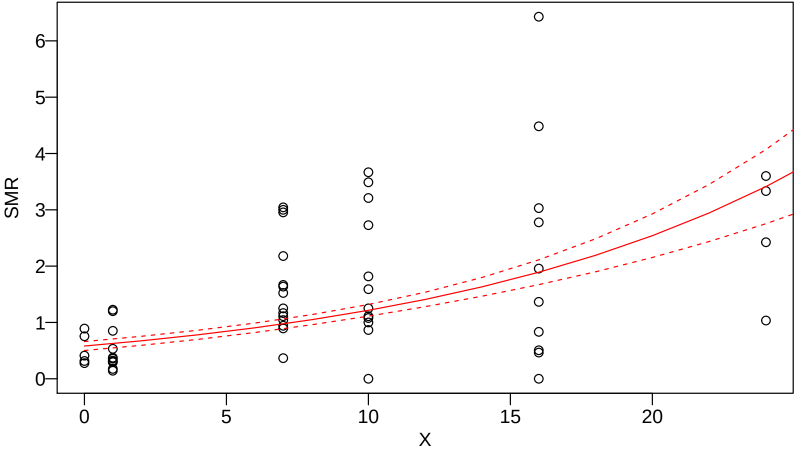
Number of **O**bserved cases as a function of the percentage working in agriculture (**X**)

```
m1 <- glm(O ~ X, poisson, offset=log(E), data=map@data)
summary(m1)
##
## Call:
## glm(formula = O ~ X, family = poisson, data = map@data, offset = log(E))
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -4.763  -1.216   0.097   1.336   4.713
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.54227    0.06952    -7.8  6.2e-15 ***
## X            0.07373    0.00596    12.4  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 380.73  on 55  degrees of freedom
## Residual deviance: 238.62  on 54  degrees of freedom
## AIC: 450.6
##
## Number of Fisher Scoring iterations: 5
```

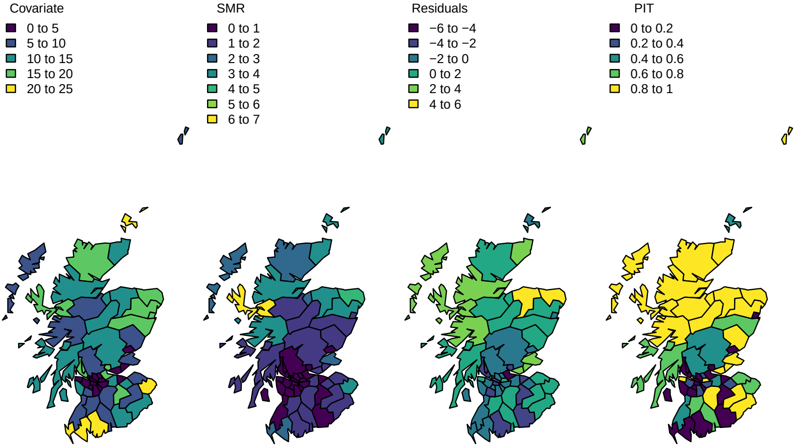
Scotland data: model 1

```
r1 <- inla(O ~ X, family='poisson', offset=log(E), data=map@data,
           control.compute=list(cpo=TRUE))
summary(r1)
##
## Call:
##    c("inla(formula = O ~ X, family = \"poisson\", data = map@data,
##    offset = log(E), ", " control.compute = list(cpo = TRUE))")
## Time used:
##     Pre = 1.47, Running = 0.0675, Post = 1.33, Total = 2.86
## Fixed effects:
##               mean    sd 0.025quant 0.5quant 0.975quant   mode kld
## (Intercept) -0.542 0.070     -0.681   -0.541     -0.407 -0.540   0
## X            0.074 0.006      0.062    0.074      0.085  0.074   0
##
## Expected number of effective parameters(stdev): 2.00(0.00)
## Number of equivalent replicates : 27.96
##
## Marginal log-Likelihood:  -234.09
## CPO and PIT are computed
##
## Posterior marginals for the linear predictor and
##  the fitted values are computed
```

The fitted covariate effect



After the covariate effect, is there something left?



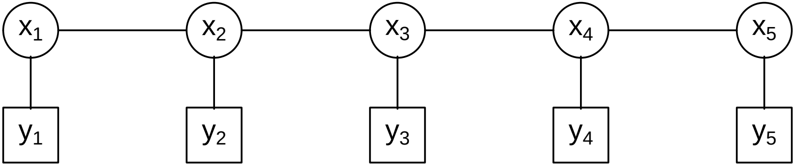
Structured random effects

Smoothed probability over time

- Temporally smooth probability of rain
  - is different for each day but similar for nearby days
    - p_i is similar to p_{i+1}
    - assume $\text{logit}(p_i) = x_i$

Smoothed probability over time

- Temporally smooth probability of rain
  - is different for each day but similar for nearby days
    - p_i is similar to p_{i+1}
    - assume $\text{logit}(p_i) = x_i$



- dependence on x
- y conditionally independent given x
  - y_i conditional on x_i is independent of y_{i-1} and of y_{i+1}

The RW1 prior

- It seems natural to borrow strength over time.
  - **x**: smoothing over time
  - ***R****andon* ***W****alk* - RW of first order: `rw1`
  - Gaussian distribuiton for the sucessive differences (**R** esparse)

$$\Delta_i = x_i - x_{i-1} \sim N(0, \tau^{-1})$$

The RW1 prior

- It seems natural to borrow strength over time.
  - **x**: smoothing over time
  - ***R****andon* ***W****alk* - RW of first order: `rw1`
  - Gaussian distribuiton for the sucessive differences (**R** esparse)

$$\Delta_i = x_i - x_{i-1} \sim N(0, \tau^{-1})$$

- The log of the (joint) distribution for **x** is

$$\log(\pi(\mathbf{x}|\tau)) \propto -\frac{\tau}{2}\sum_{i=2}^{n}(x_i - x_{i-1})^2 = -\frac{\tau}{2}\mathbf{x}'\mathbf{R}\mathbf{x},$$

where

$$\mathbf{R} = \begin{bmatrix} 1 & -1 & & & & & & \\ -1 & 2 & -1 & & & & & \\ & -1 & 2 & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & -1 & 2 & -1 & \\ & & & & & & -1 & 2 & -1 \\ & & & & & & & -1 & 1 \end{bmatrix}$$

The cyclic RW1

- It is also reasonable to assume that 1st of January is similar to December, 31: cyclic random walk
- A random walk of first order (CRW1) is defined as:

$$\pi(\mathbf{x}|\theta) \propto \exp\left\{-\frac{\theta}{2}\left[(x_1 - x_n)^2 + \sum_{i=2}^{n}(x_i - x_{i-1})^2\right]\right\} = \exp\left\{-\frac{\theta}{2}\mathbf{x}^T\mathbf{R}\mathbf{x}\right.$$

where, now,

$$\mathbf{R} = \begin{bmatrix} 2 & -1 & & & & & & -1 \\ -1 & 2 & -1 & & & & & \\ & -1 & 2 & -1 & & & & \\ & & & & \ddots & & & \\ & & & & -1 & 2 & -1 & \\ & & & & & -1 & 2 & -1 \\ -1 & & & & & & -1 & 2 \end{bmatrix}$$

- The definition of a random walk of second order (CRW2) is analogous.

Tokyo example: the model

- y_i assume values 0, 1 or 2, for $i = 1, ..., n$
  - assuming conditional independence, thus

$$y_i | p_i \sim \text{Binomial}(n_i, p_i)$$

  - link function (logit)

$$p_i = 1/(1 + \exp(-x_i))$$

Tokyo example: the model

- y_i assume values 0, 1 or 2, for $i = 1, ..., n$
  - assuming conditional independence, thus

$$y_i | p_i \sim \mathrm{Binomial}(n_i, p_i)$$

  - link function (logit)

$$p_i = 1/(1 + \exp(-x_i))$$

- **x** is a *Gaussian Markov Random Field* - GMRF, Rue and Held (2005)

Tokyo example: the model

- y_i assume values 0, 1 or 2, for $i = 1, ..., n$
  - assuming conditional independence, thus

$$y_i | p_i \sim \text{Binomial}(n_i, p_i)$$

  - link function (logit)

$$p_i = 1/(1 + \exp(-x_i))$$

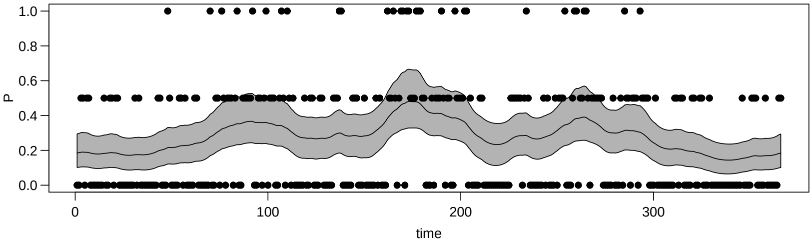
- **x** is a *Gaussian Markov Random Field* - GMRF, Rue and Held (2005)
- τ: local precision parameter

Model fit in INLA

$$
\begin{aligned}
y_i|x_i &\sim \text{Binomial}(2, p_i) && \rightarrow \text{likelihood} \\
\mathbf{x}|\tau &\sim N(\mathbf{0}, (\tau\mathbf{R})^{-}) && \rightarrow \text{latent field, GMRF} \\
\tau &\sim p(\tau) && \rightarrow \text{prior distribution}
\end{aligned}
$$

```r
head(Tokyo, 5)
##   y n time   P
## 1 0 2    1 0.0
## 2 0 2    2 0.0
## 3 1 2    3 0.5
## 4 1 2    4 0.5
## 5 0 2    5 0.0
model <- y ~ f(time, model='rw1', cyclic=TRUE)
result <- inla(model, family='binomial',
               data=Tokyo, Ntrials=n,
               control.compute=list(cpo=TRUE))
```

Result for the time series



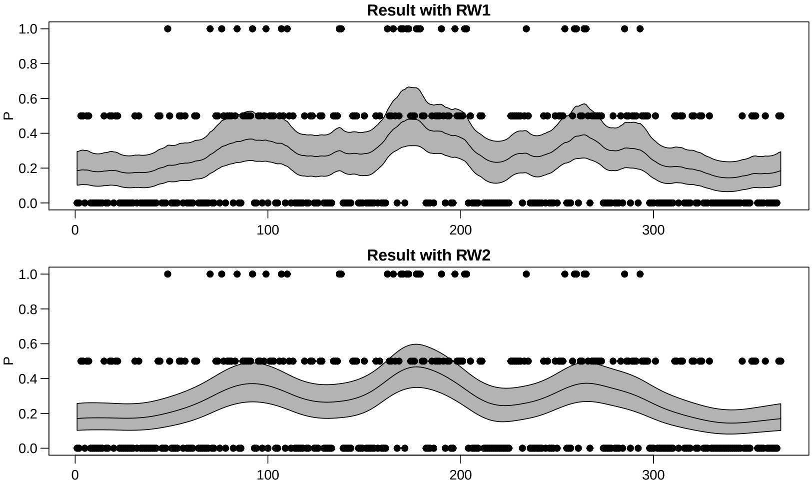
Smoothing more

Gaussian distribution for the second order differences (`rw2`)

$$\Delta_i^2 = x_i - 2x_{i-1} + x_{i-2} \sim N(0, \tau^{-1})$$

```
model2 <- y ~ f(time, model='rw2', cyclic=TRUE)
result2 <- inla(model2, family='binomial',
                data=Tokyo, Ntrials=n,
                control.compute=list(cpo=TRUE))
```

Both results for the time series



Smooth areal dependent risk

$$y_i \sim \text{Poisson}(E_i r_i)$$

$$\log(r_i) = \alpha + \beta X_i + s_i$$

where s_i may be spatial smooth

$s_i | s_j$, j the index for the neighbours of i

The Besag's model: randon walk over areas, Besag (1974)

$$\pi(x_i|\mathbf{x}_{-i}, \tau) \sim N(\frac{1}{n_i}\sum_{j\sim i} x_j, \frac{1}{n_i \tau})$$

where $j \sim i$ means j *neighbour of* i. This gives:

The Besag's model: randon walk over areas, Besag (1974)

$$\pi(x_i|\mathbf{x}_{-i},\tau) \sim N(\frac{1}{n_i}\sum_{j\sim i} x_j, \frac{1}{n_i\tau})$$

where $j \sim i$ means j *neighbour of* i. This gives:

$$\pi(\mathbf{x}|\tau) \quad \propto \quad \tau^{(n-1)/2} \exp\left(-\frac{\tau}{2}\sum_{i}^{n}(x_i - \frac{1}{n_i}\sum_{j\sim i} x_j)^2\right)$$

The Besag's model: randon walk over areas, Besag (1974)

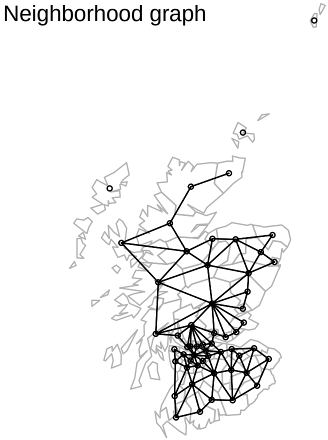
$$\pi(x_i|\mathbf{x}_{-i},\tau) \sim N(\frac{1}{n_i}\sum_{j\sim i} x_j, \frac{1}{n_i\tau})$$

where $j \sim i$ means j *neighbour of* i. This gives:

$$\begin{aligned}\pi(\mathbf{x}|\tau) &\propto \tau^{(n-1)/2}\exp\left(-\frac{\tau}{2}\sum_{i}^{n}(x_i - \frac{1}{n_i}\sum_{j\sim i} x_j)^2\right) \\ &= \tau^{(n-1)/2}\exp\left(-\frac{\tau}{2}\sum_{j\sim i}^{n}(x_i - x_j)^2\right) = \tau^{(n-1)/2}\exp\left(-\frac{\tau}{2}\mathbf{x}^T\mathbf{R}\mathbf{x}\right)\end{aligned}$$

$$\mathbf{R}_{ij} = \begin{cases} n_i & \text{if } i = j \\ -1 & \text{if } j \sim i \\ 0 & \text{otherwise} \end{cases}.$$

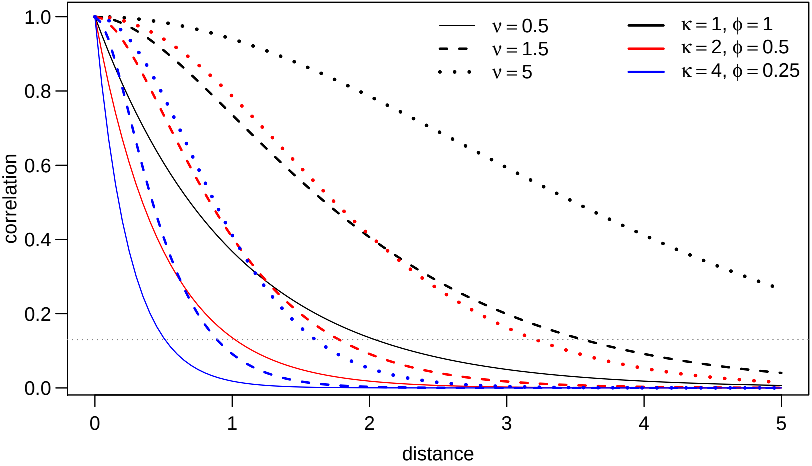
The Scotland graph



The SPDE model approach

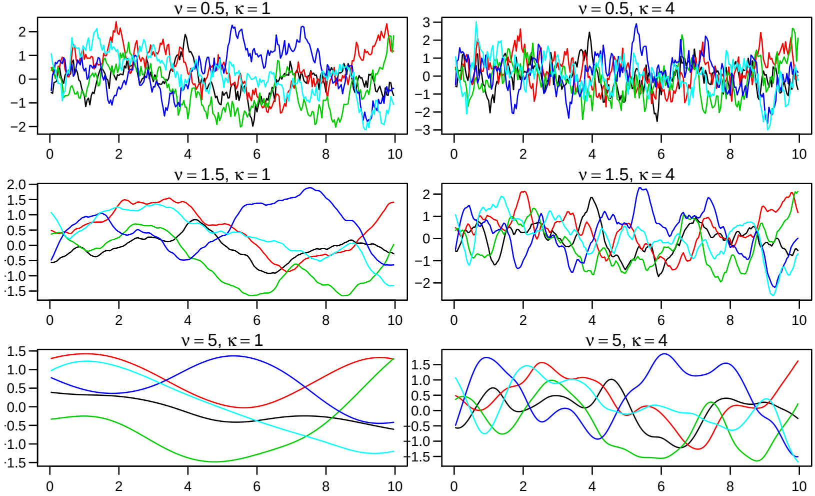
The Mat'ern covariance

$$\Sigma_{ij} = \sigma_x^2 \frac{2^{1-\nu} K_\nu(\kappa\|\mathbf{s}_i - \mathbf{s}_j\|)}{\Gamma(\nu)(\kappa\|\mathbf{s}_i - \mathbf{s}_j\|)^{-\nu}},\ \kappa = 1/\phi$$



corr $((8\nu)^{1/2}/\kappa) \approx 0.13$

Simulations, 1D, $\sigma_x^2 = 1$



The Stochastic Partial Differential Approach - SPDE

- Fields with Matérn covariance are solutions to (SPDE), Whittle (1954) and Lindgren, Rue, and Lindström (2011)

$$(\kappa^2 - \Delta)^{\alpha/2}\xi(\mathbf{s}) = \tau\mathcal{W}(\mathbf{s})$$

  - $\kappa > 0$: scale parameter
  - $\alpha = \nu + d/2$: smoothness
  - Δ is the Laplacian

$$\Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial s_i^2}$$

The Stochastic Partial Differential Approach - SPDE

- Fields with Matérn covariance are solutions to (SPDE), Whittle (1954) and Lindgren, Rue, and Lindström (2011)

$$(\kappa^2 - \Delta)^{\alpha/2}\xi(\mathbf{s}) = \tau\mathcal{W}(\mathbf{s})$$

  - $\kappa > 0$: scale parameter
  - $\alpha = \nu + d/2$: smoothness
  - Δ is the Laplacian

$$\Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial s_i^2}$$

  - When $d = 2$
    - $\alpha = 1$: CAR model
    - $\alpha = 2$: SAR model

Regular grid, $d = 2$

- $\alpha = 1$: $\mathbf{Q}_{1,\kappa} = \mathbf{K}_\kappa = \kappa^2 \mathbf{C} + \mathbf{G}$
- $\mathbf{C} = \mathbf{I}$, $\mathbf{G} =$ Laplacian (4 neighbours)
  - Laplacian-local pattern:

$$\begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix}$$

  - $\mathbf{Q}_{1,\kappa}$-local pattern

$$\begin{bmatrix} & -1 & \\ -1 & 4 + \kappa^2 & -1 \\ & -1 & \end{bmatrix}$$

Regular grid, $d = 2$

- $\alpha = 1$: $\mathbf{Q}_{1,\kappa} = \mathbf{K}_\kappa = \kappa^2\mathbf{C} + \mathbf{G}$
- $\mathbf{C} = \mathbf{I}$, $\mathbf{G}$ = Laplacian (4 neighbours)
  - Laplacian-local pattern:

$$\begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix}$$

  - $\mathbf{Q}_{1,\kappa}$-local pattern

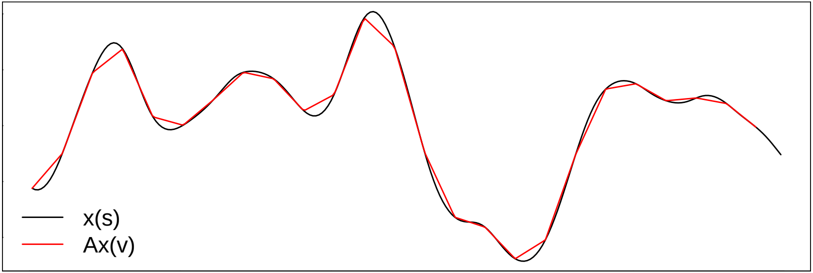
$$\begin{bmatrix} & -1 & \\ -1 & 4+\kappa^2 & -1 \\ & -1 & \end{bmatrix}$$

- κ is a scale parameter
  - $\rightarrow$ Sparse precision $\mathbf{Q}$ {**!!!**}
  - remember: $(\kappa^2 - \Delta)^{\alpha/2}\xi(\mathbf{s}) = \tau\mathcal{W}(\mathbf{s})$
  - $\rightarrow \quad (\mathbf{Q}_{1,\kappa})^{1/2}\xi$ = independent noise
  - 'effective' range (0.139) $\approx \sqrt{8\nu/\kappa}$

Important fact: role of α

- Bigger $\alpha \rightarrow \mathbf{Q}$ less sparse $\rightarrow$ smoother
  - $\alpha = 1$: $\mathbf{Q}_{1,\kappa} = \mathbf{K}_\kappa = \kappa^2\mathbf{C} + \mathbf{G}$
  - $\alpha = 2$: $\mathbf{Q}_{2,\kappa} = \mathbf{K}_\kappa\mathbf{C}^{-1}\mathbf{K}_\kappa$
  - $\alpha = 3, 4, ...$: $\mathbf{Q}_{\alpha,\kappa} = \mathbf{K}_\kappa\mathbf{C}^{-1}\mathbf{Q}_{\alpha-2,\kappa}\mathbf{C}^{-1}\mathbf{K}_\kappa$

Irregular grid → Finite Element Method - FEM → mesh



- $\xi(\mathbf{s}) \approx \sum_{k=1}^{m} \psi_k(\mathbf{s}) w_k = \mathbf{A}\xi(\mathbf{v}$,
  - ψ_k: basis functions,
  - w_k: weights

Basis functions plot

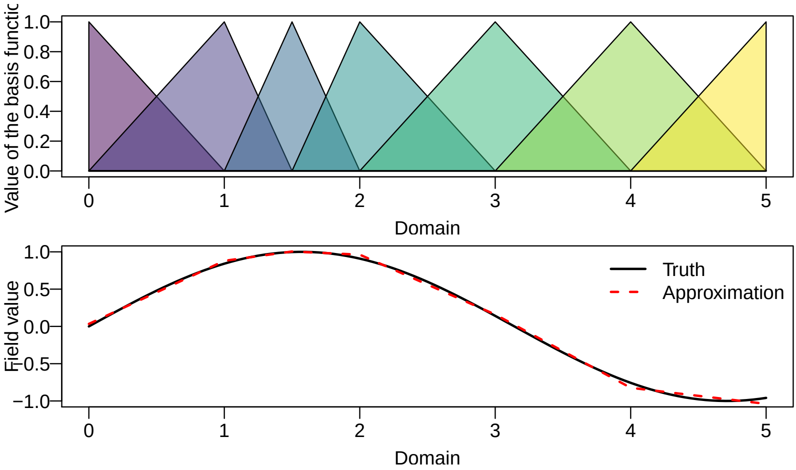


Figure 2: One dimensional approximation illustration. The one dimensional piece-wise linear basis functions (top). A function and its approximation (bottom).

References

References

Besag, J. 1974. "Spatial Interaction and the Statistical Analysis of Lattice Systems." *JRSS-B* 36 (2): 192–236.

Lindgren, F., H. Rue, and J. Lindström. 2011. "An Explicit Link Between Gaussian Fields and Gaussian Markov Random Fields: The Stochastic Partial Differential Equation Approach (with Discussion)." *JRSS-B* 73 (4): 423–98.

Rue, H., and L. Held. 2005. *Gaussian Markov Random Fields: Theory and Applications*. Monographs on Statistics & Applied Probability. Boca Raton: Chapman; Hall.

Whittle, P. 1954. "On Stationary Processes in the Plane." *Biometrika* 41 (3/4): 434–49.