Model

\[ Y = X\beta + B(Y - X\beta) + E \]

with

\[ Y = (Y_1, \ldots, Y_{49})' \]: the crime statistics per district

\( \beta \) has three components: an offset value, ‘house value’ and ‘income’

\( X \) is a \( 49 \times 3 \) matrix whose first column contains 1s, the second column the house values, and the third the income values.

\( E \) is spatially correlated noise with mean zero and covariance matrix \( \sigma^2(I - B) \) for \( B = \rho N \).

The \( N \) were obtained in the homework for Week 5.
There are five parameters to estimate: the three components of $\beta$, $\rho$ and $\sigma^2$. The fitted values are $X\hat{\beta} + \hat{\rho}N(Y - X\hat{\beta})$. Subtracting from the observations gives the residuals.

The model can be fitted as follows. (The style indicates we use the binary neighbourhood coding).

```r
library("spdep")
columbus.listw <- nb2listw( columbus.nb, style="B")
car.out <- spautolm( formula= CRIME ~ HOVAL + INC, 
  data=columbus.poly, listw=columbus.listw, family="CAR" )
columbus.poly$fitted.car <- fitted( car.out )
```
> print(car.out)

Call:
spautolm(formula = CRIME ~ HOVAL + INC, data = columbus.poly,
        listw = columbus.listw, family = "CAR")

Coefficients:
(Intercept)     HOVAL      INC     lambda
 54.3139189  -0.2821969 -0.9882862  0.1589004

Log likelihood:  -182.2198

Less crime in districts with high HOVAL and INC. Note that \( \rho \) is called \( \lambda \) here. As for \( \hat{\sigma}^2 \),

> car.out$fit$s2
[1] 87.65356
library("classInt")
library("RColorBrewer")
fitted.colours <- classIntervals(var=columbus.poly$fitted.car,
                                 n=9, style="fixed", fixedBreaks=breaks)
plot(columbus.poly,
     col= findColours(fitted.colours, colour.palette) )

Note the smoothing effect.
Ohio crime data: Residuals.

> plot(car.out$fit$residuals)
Image segmentation

Let $x \in L^T$ denote the target labelled image and $y \in R^T$ the observed one.

**Forward model**

$$f(y|x) = \prod_{i \in T} g(y_i|x_i)$$

where $g(\cdot|x_i)$ is the p.d.f. of the observations in label class $x_i$, e.g.

$$g(y_i|x_i) \propto \exp\left(-\frac{1}{2\sigma^2}(y_i - x_i)^2\right).$$

**Maximum likelihood estimator**

$$\hat{x}_i = \arg\max \\{g(y_i|x_i) : x_i \in L\}, \quad i \in T.$$
Bayesian approach

with prior distribution $\pi_X(x)$ to penalise rough labellings.

Posterior

$$f(x|y) \propto f(y|x)\pi_X(x) = \pi_X(x) \prod_{i \in T} g(y_i|x_i).$$

MAP estimator

$$\tilde{x} = \arg\max \{ f(y|x)\pi_X(x) : x \in L^T \}$$
$$= \arg\max \{ \log f(y|x) + \log \pi_X(x) : x \in L^T \}. $$
Potts model

\[ \pi_X(x) \propto \exp \left[ -\theta \sum_{i \sim j; i<j} 1\{x_i \neq x_j\} \right] \]

for \( \theta > 0 \).

Greedy iterative pointwise optimisation:

\[ \tilde{x}_i = \arg\max \ g(y_i|x_i) \pi_i(x_i|x_{T\setminus i}) \]

starting in e.g. the maximum likelihood estimator \( \hat{x} \).
Segmentation of a noisy image of a cat. From left to right: Truth, distortion by white noise, MLE and MAP classifiers.
Disease mapping

\[ \pi_i(y_i|y_{T\setminus i}) = \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!}, \quad y_i \in \{0, 1, \ldots\}, \]

with

\[ \log \mu_i = \alpha_i + \theta \sum_{j \sim i} y_j. \]

Since \( L = \{0, 1, \ldots\} \), we must verify that any joint distribution is well-defined.

**Homework:** \( \theta \leq 0 \). Also the logarithmic transform does not scale w.r.t. the size of the areal units.
Mixture model

Set, for $\alpha = X\beta$,

$$\mu_i = c_i e^{\alpha_i} \Lambda Z_i$$

where $c_i$ is a base rate of expected counts, $\Lambda Z_i$ is the area-specific relative risk.

The $Z_i$ assign $i$ to one of $k$ mixture components $\lambda_1, \ldots, \lambda_k \geq 0$. To achieve spatial coherence, assume that $(Z_i)_{i \in T}$ follow a Potts model.

Provided the covariates do not fluctuate too wildly within the areal unit, $\mu_i$ scales appropriately with size.
Inference for $\lambda_Z$ as well as the model parameters $\beta$, $\theta$ and $\lambda_1, \ldots, \lambda_k$.

Forward model

$$f(y|z; \alpha, \lambda_1, \ldots, \lambda_k) = \prod_{i \in T} g(y_i|z_i; \alpha_i, \lambda_{z_i})$$

with $g(\cdot|z_i; \alpha_i, \lambda_{z_i})$ the p.d.f. of a Poisson with mean $\mu_i = c_i e^{\alpha_i \lambda_{z_i}}$.

Posterior

$$f(z|y; \alpha, \theta, \lambda_1, \ldots, \lambda_k) \propto f(y|z; \alpha, \lambda_1, \ldots, \lambda_k) \pi_Z(z; \theta)$$

or

$$f(y|z; \alpha, \lambda_1, \ldots, \lambda_k) \pi_Z(z; \theta, k)p(\lambda_1, \ldots, \lambda_k|k)p(\theta)p(k).$$
Posterior mean (left) and posterior distribution (right) of risk for Larynx cancer in France (1986–1993).
Synthesis

A spatial process of interest cannot be observed directly but only through other random variables and the model contains unknown parameters. I.e. the joint distribution takes the form

\[
\text{forward model}[\text{data} \mid \text{process}, \text{parameters}] \times \text{prior}[\text{process} \mid \text{parameters}],
\]

and inference is based on the posterior distribution of the process and/or the parameters conditional on the observations.