Statistical inference for clustered point patterns

Rasmus Waagepetersen
Department of Mathematical Sciences
Aalborg University

September 7, 2011

Outline:

Background in tropical forest ecology

Intensity function and second order summary statistics

Poisson and Cox processes

Estimation

Optimal first-order estimating equations - quasi-likelihood

Decomposition of variance

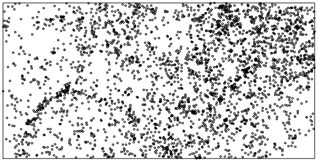
Background: Tropical rain forest ecology

Fundamental questions: which factors influence the spatial distribution of rain forest trees and what is the reason for the high biodiversity of rain forests ?

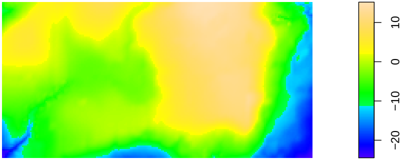
Key factors:

- environment: topography, soil composition,...
- seed dispersal limitation: by wind, birds or mammals...

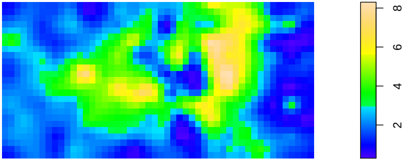
Data example: *Capparis Frondosa*



- *observation window W* $= 1000 \text{ m} \times 500 \text{ m}$
- seed dispersal$\Rightarrow$ *clustering*
- environment $\Rightarrow$ *inhomogeneity*



Elevation



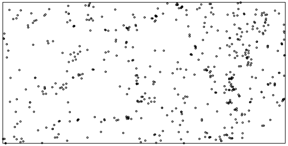
Potassium content in soil.

Quantify dependence on environmental variables taking into account clustering due to e.g. seed dispersal.

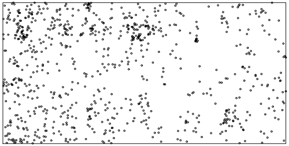
Example: modes of seed dispersal and clustering

Three species with different modes of seed dispersal:

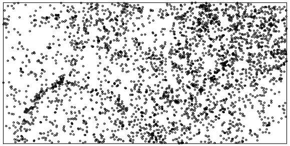
Acalypha Diversifolia explosive capsules



Loncocharpus Heptaphyllus wind



Capparis Frondosa bird/mammal



Quantify how much of the spatial variation is due to respectively environment and seed dispersal ?

Differences between species ?

Approach: Cox process model for joint effects of environment and seed dispersal.

Outline:

Background in tropical forest ecology

Intensity function and second order summary statistics

Poisson and Cox processes

Estimation

Optimal first-order estimating equations - quasi-likelihood

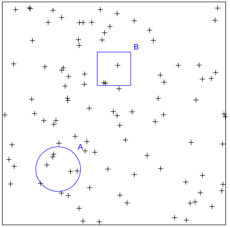
Decomposition of variance

Mean and covariances of counts for spatial point process

Point process $\mathbf{X}$: random point pattern.

For A subset of the plane, count $N(A)$ is number of points in A.

$$\mathbb{E}N(A) = \int_A \rho(u)\mathrm{d}u$$



$\rho(\cdot)$: intensity function.

$$\mathbb{E}[N(A)N(B)] = \int_{A\cap B} \rho(u)\mathrm{d}u + \int_A \int_B \rho^{(2)}(u,v)\mathrm{d}u\mathrm{d}v$$

$\rho^{(2)}(u,v)$: second order product density

Infinitesimal interpretation

Very small A and $B \Rightarrow N(A)$ and $N(B)$ binary:

$$\mathbb{E}N(A) \approx P(\mathbf{X} \text{ has a point in A}) \approx \rho(u)|A|, \quad u \in A$$

$$\mathbb{E}N(A)N(B) \approx P(\mathbf{X} \text{ has a point in each of } A \text{ and } B)$$
$$\approx \rho^{(2)}(u,v)|A||B| \quad u \in A,\ v \in B$$

Infinitesimal interpretation

Very small A and $B \Rightarrow N(A)$ and $N(B)$ binary:

$$\mathbb{E}N(A) \approx P(\mathbf{X} \text{ has a point in A}) \approx \rho(u)|A|, \quad u \in A$$

$$\mathbb{E}N(A)N(B) \approx P(\mathbf{X} \text{ has a point in each of } A \text{ and } B) \approx \rho^{(2)}(u,v)|A||B| \quad u \in A,\ v \in B$$

Pair correlation

$$g(u,v) = \frac{\rho^{(2)}(u,v)}{\rho(u)\rho(v)} = \frac{P(\mathbf{X} \text{ has a point in each of } A \text{ and } B)}{P(\mathbf{X} \text{ has a point in A})P(\mathbf{X} \text{ has a point in B})}$$

Infinitesimal interpretation

Very small A and $B \Rightarrow N(A)$ and $N(B)$ binary:

$$\mathbb{E}N(A) \approx P(\mathbf{X} \text{ has a point in A}) \approx \rho(u)|A|, \quad u \in A$$

$$\mathbb{E}N(A)N(B) \approx P(\mathbf{X} \text{ has a point in each of } A \text{ and } B)$$
$$\approx \rho^{(2)}(u,v)|A||B| \quad u \in A,\ v \in B$$

Pair correlation

$$g(u,v) = \frac{\rho^{(2)}(u,v)}{\rho(u)\rho(v)} = \frac{P(\mathbf{X} \text{ has a point in each of } A \text{ and } B)}{P(\mathbf{X} \text{ has a point in A})P(\mathbf{X} \text{ has a point in B})}$$

$$= 1 \text{ if independence (Poisson process)}$$

K-function

$$K(t) = \int_{\|h\| \le t} g(h) \mathrm{d}h$$

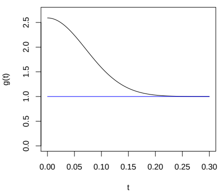
(provided $g(u, v) = g(u - v)$ i.e. $\mathbf{X}$ second-order reweighted stationary)

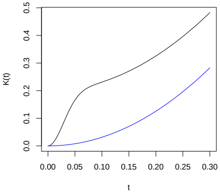
K-function

$$K(t) = \int_{\|h\| \le t} g(h) \mathrm{d}h$$

(provided $g(u, v) = g(u - v)$ i.e. $\mathbf{X}$ second-order reweighted stationary)

Examples of pair correlation and K-functions:



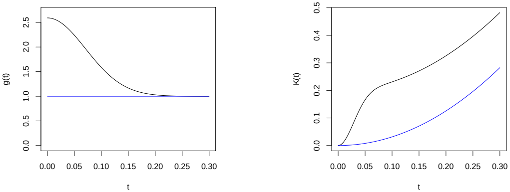


K-function

$$K(t) = \int_{\|h\| \le t} g(h) \mathrm{d}h$$

(provided $g(u, v) = g(u - v)$ i.e. $\mathbf{X}$ second-order reweighted stationary)

Examples of pair correlation and K-functions:



Unbiased estimate of K-function (W observation window):

$$\hat{K}_{\beta}(t) = \sum_{u,v \in \mathbf{X} \cap W} \frac{1[0 < \|u - v\| \le t]}{\rho(u)\rho(v)} e_{u,v}$$

($e_{u,v}$ edge correction factor)

Outline:

Background in tropical forest ecology

Intensity function and second order summary statistics

Poisson and Cox processes

Estimation

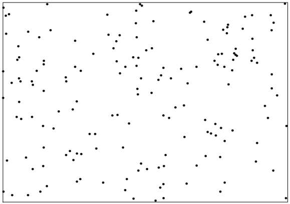
Optimal first-order estimating equations - quasi-likelihood

Decomposition of variance

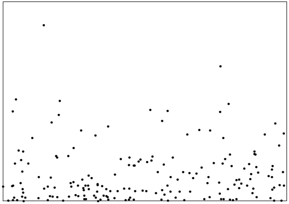
The Poisson process

$\mathbf{X}$ is a Poisson process with intensity function $\rho(\cdot)$ if for any bounded region B:

1. $N(B)$ is Poisson distributed with mean $\mu(B) = \int_B \rho(u)\mathrm{d}u$
2. Given $N(B) = n$, the n points are independent and identically distributed with density proportional to intensity function $\rho(\cdot)$.



Homogeneous: $\rho = 150/0.7$



Inhomogeneous: $\rho(x, y) \propto \mathrm{e}^{-10.6y}$

Back to rain forest: parametric models for intensity and pair correlation

Study influence of covariates

$$Z(u) = (Z_1(u), \ldots, Z_p(u))$$

using log-linear model for intensity function:

$$\rho(u; \beta) = \exp[\beta Z(u)^\mathsf{T}]$$

where

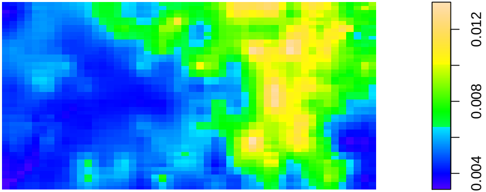
$$\beta Z(u)^\mathsf{T} = \beta_1 Z_1(u) + \beta_2 Z_2(u) + \ldots + \beta_p Z_p(u)$$

Capparis Frondosa and Poisson process ?

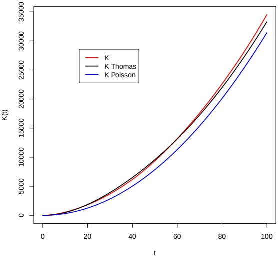
Fit model with covariates elevation and Potassium.

Fitted intensity function

$\rho(u;\hat{\beta}) = \exp(\hat{\beta}_0 + \hat{\beta}_1 \mathrm{Elev}(u) + \hat{\beta}_2 \mathrm{K}(u))$:

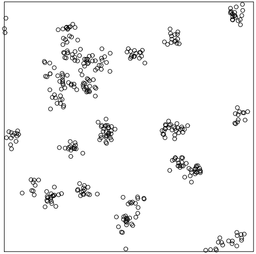


Estimated K-function and $K(t) = \pi t^2$-function for Poisson process:



Not Poisson process - aggregation due to unobserved factors (e.g. seed dispersal)

Cluster process: Inhomogeneous Thomas process (W, 2007)



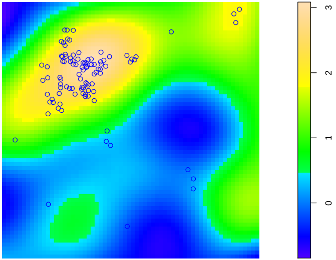
Parents stationary Poisson point process intensity κ

Offspring distributed around mothers according to Gaussian density with standard deviation ω

Inhomogeneity: offspring survive according to probability

$$p(u) \propto \exp(Z(u)\beta^{\mathsf{T}})$$

depending on covariates (independent thinning).



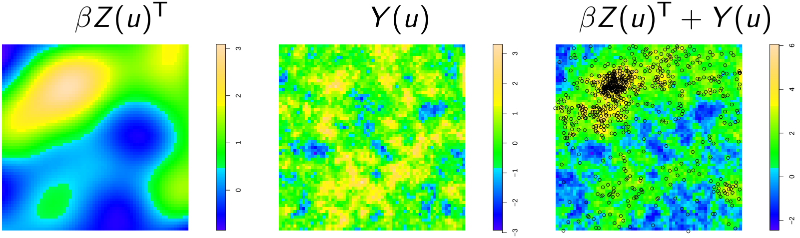
Cox processes

X is a *Cox process* driven by the random intensity function Λ if, conditional on $\Lambda = \lambda$, **X** is a Poisson process with intensity function λ.

Example: log Gaussian Cox process (Møller, Syversveen, W, 1998)

$$\log \Lambda(u) = \beta Z(u)^{\mathsf{T}} + Y(u)$$

where $\{Y(u)\}$ Gaussian random field.



Shot-noise Cox process

$$\Lambda(u) = \sum_{v \in C} \gamma_v k(u - v)$$

where

- C homogeneous Poisson with intensity κ
- $k(\cdot)$ probability density.
- γ_v *iid* positive random variables independent of C

Shot-noise Cox process

$$\Lambda(u) = \sum_{v \in C} \gamma_v k(u - v)$$

where

- C homogeneous Poisson with intensity κ
- $k(\cdot)$ probability density.
- γ_v *iid* positive random variables independent of C

NB: equivalent to cluster process with parents C, random cluster size γ_v and dispersal density f.

Shot-noise Cox process

$$\Lambda(u) = \sum_{v \in C} \gamma_v k(u - v)$$

where

- C homogeneous Poisson with intensity κ
- $k(\cdot)$ probability density.
- γ_v *iid* positive random variables independent of C

NB: equivalent to cluster process with parents C, random cluster size γ_v and dispersal density f.

Inhomogeneous shot-noise:

$$\Lambda(u) = \exp(\beta Z(u)^{\mathsf{T}}) \sum_{v \in C} \gamma_v k(u - v)$$

Inhomogeneous Thomas: inhomogeneous shot-noise with Gaussian $k(\cdot)$.

Moments for Cox processes

Intensity function

$$\rho(u) = \mathbb{E}\Lambda(u)$$

Second-order product density

$$\rho^{(2)}(u,v) = \mathbb{E}\Lambda(u)\Lambda(v) = \mathbb{C}\mathrm{ov}[\Lambda(u),\Lambda(v)] + \rho(u)\rho(v)$$

$$\begin{aligned}
\mathbb{C}\mathrm{ov}[N(A),N(B)] &= \int_{A\cap B} \mathbb{E}\Lambda(u)\mathrm{d}u + \int_A \int_B \mathbb{C}\mathrm{ov}[\Lambda(u),\Lambda(v)]\mathrm{d}u\mathrm{d}v \\
&= \int_{A\cap B} \rho(u)\mathrm{d}u + \int_A \int_B \rho(u)\rho(v)[g(u,v)-1]\mathrm{d}u\mathrm{d}v \\
&= \text{ Poisson variance } + \text{ extra variance due to } \Lambda
\end{aligned}$$

Log-linear model

$$\Lambda(u) = \Lambda_0(u) \exp[\beta Z(u)^\mathsf{T}]$$

where Λ_0 stationary non-negative reference process (both log Gaussian Cox process and inhom. shot-noise of this form).

Log-linear intensity (assume $\mathbb{E}\Lambda_0(u) = 1$)

$$\rho(u) = \mathbb{E}\Lambda(u) = \exp[\beta Z(u)^\mathsf{T}]$$

Pair correlation function ($\mathbb{E}\Lambda_0(u) = 1$):

$$g(h) = 1 + c_0(h) \quad c_0(h) = \mathbb{C}\mathrm{ov}[\Lambda_0(u), \Lambda_0(u+h)]$$

Specific models for $c_0(u - v) = \mathbb{C}\mathrm{ov}[\Lambda_0(u), \Lambda_0(v)]$

Log-Gaussian:

$$\Lambda_0(u) = \exp[Y(u)]$$

where Y Gaussian field.

Covariance (Laplace transform):

$$c_0(h) = \exp[\mathbb{C}\mathrm{ov}(Y(u), Y(u+h))] - 1$$

Specific models for $c_0(u-v) = \mathbb{C}\mathrm{ov}[\Lambda_0(u), \Lambda_0(v)]$

Log-Gaussian:

$$\Lambda_0(u) = \exp[Y(u)]$$

where Y Gaussian field.

Covariance (Laplace transform):

$$c_0(h) = \exp[\mathbb{C}\mathrm{ov}(Y(u), Y(u+h))] - 1$$

Shot-noise:

$$\Lambda_0(u) = \sum_{v \in C} \gamma_v k(u-v)$$

Covariance (convolution):

$$c_0(u-v) = \kappa\alpha^2 \int_{\mathbb{R}^2} k(u)k(u+h)\mathrm{d}u$$

$(\alpha = \mathbb{E}\gamma_v)$

Bessel shot-noise/Matérn covariance

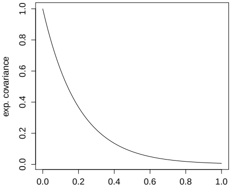
Suppose kernel $k(\cdot)$ given by variance-gamma density (Bessel density).

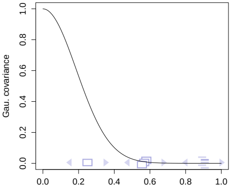
Y variance-gamma if $Y = \sqrt{W}Z$ where $W \sim \Gamma$ and $Z \sim N_p(0, I)$ $\Rightarrow$ closed under convolution.

Then Matérn covariance function:

$$c_0(h) = \sigma_0^2 \frac{(\|h\|/\eta)^\nu K_\nu(\|h\|/\eta)}{2^{\nu-1}\Gamma(\nu)}$$

$\nu = 1/2$: exponential model ‘$\nu = \infty$’:‘Gaussian’ (mod. Thomas)





Outline:

Background in tropical forest ecology

Intensity function and second order summary statistics

Poisson and Cox processes

Estimation

Optimal first-order estimating equations - quasi-likelihood

Decomposition of variance

Maximum likelihood estimation for Cox processes

Suppose $\mathbf{x}$ observed point pattern (realization of $\mathbf{X}$ inside observation window W). Likelihood (probability density) given Λ (Poisson process):

$$\prod_{u \in \mathbf{x}} \Lambda(u) \exp\left[-\int_W \Lambda(u)\mathrm{d}u\right]$$

Likelihood for Cox process (integrate out unobserved Λ):

$$f_\theta(\mathbf{x}) = \mathbb{E}_\theta\left[\prod_{u \in \mathbf{x}} \Lambda(u) \exp\left[-\int_W \Lambda(u)\mathrm{d}u\right]\right]$$

Problem for Monte Carlo approximation: $\Lambda = \{\Lambda(u)\}_{u \in W}$ infinitely dimensional quantity.

Estimation of regression parameters

For log-linear model,

$$\Lambda(u) = \exp(\beta Z(u)^{\mathsf{T}})\Lambda_0(u)$$

intensity function is known:

$$\rho_\beta(u) = \exp(\beta Z(u)^{\mathsf{T}})$$

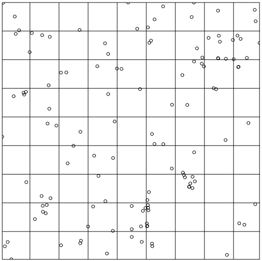
Poisson "likelihood"

$$\Big[\prod_{u \in \mathbf{X} \cap W} \rho_\beta(u)\Big] \exp\Big(-\int_W \rho_\beta(u)\,\mathrm{d}u\Big)$$

may be viewed as a composite likelihood for estimating β.

Composite likelihood obtained from binary random field

Disjoint subdivision $W = \cup_{i=1}^{m} C_i$ in 'cells' C_i.

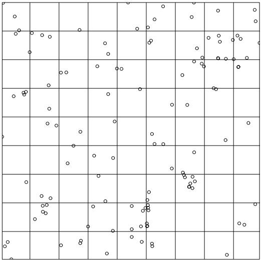


Random count variables:
$N_i = N(C_i) = \#\mathbf{X} \cap C_i$ number of points in C_i.

$X_i = 1[N_i > 0]$ binary random variable.
$P_\beta(X_i = 1) = \rho_\beta(u_i)|C_i|$.

Composite likelihood obtained from binary random field

Disjoint subdivision $W = \cup_{i=1}^m C_i$ in 'cells' C_i.



Random count variables:
$N_i = N(C_i) = \#\mathbf{X} \cap C_i$ number of points in C_i.

$X_i = 1[N_i > 0]$ binary random variable.
$P_\beta(X_i = 1) = \rho_\beta(u_i)|C_i|$.

Bernouilli composite likelihood

$$\prod_{i=1}^m P_\beta(X_i = 1)^{X_i}(1 - P_\beta(X_i = 1))^{1-X_i} \equiv \prod_{i=1}^m \rho_\beta(u_i)^{X_i}(1 - \rho_\beta(u_i)|C_i|)^{1-X_i}$$

has limit ($|C_i| \to 0$)

$$\mathrm{CL}(\beta) = \Big[\prod_{u \in \mathbf{X}\cap W} \rho_\beta(u)\Big] \exp\Big(-\int_W \rho_\beta(u)\mathrm{d}u\Big)$$

Estimation of pair correlation function

Suppose parametric model $g(\cdot;\psi)$ for pair correlation.

Some options:

1. minimum contrast estimation based on so-called K-function.
2. second-order composite likelihood: composite likelihood based on indicators for joint occurrence of points in pairs of cells:

$$X_{ij} = 1[N_i > 0 \text{ and } N_j > 0]$$

$$P_{\beta,\psi}(X_{ij} = 1) = \rho_\beta(u_i)\rho_\beta(v_j)g(u_i - u_j;\psi)$$

Minimum contrast estimation for ψ

Computationally easy alternative if $\mathbf{X}$ second-order reweighted stationary so that K-function well-defined.

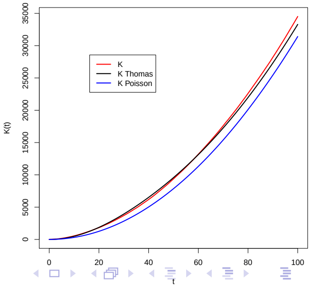
Estimate of K-function:

$$\hat{K}_\beta(t) = \sum_{u,v \in \mathbf{X} \cap W} \frac{1[0 < \|u - v\| \le t]}{\rho(u;\beta)\rho(v;\beta)} e_{u,v}$$

Unbiased if β 'true' regression parameter.

Minimum contrast estimation: minimize squared distance between theoretical K and $\hat{K}$:

$$\hat{\psi} = \underset{\psi}{\operatorname{argmin}} \int_0^r \left(\hat{K}_{\hat{\beta}}(t) - K(t;\psi)\right)^2 \mathrm{d}t$$



Second-order composite likelihood

Second-order composite likelihood (given $\hat{\beta}$):

$$\mathrm{CL}_2(\psi|\hat{\beta}) = \sum^{\neq}_{\substack{u,v \in \mathbf{X} \cap W \\ \|u-v\| \leq R}} \log \rho^{(2)}(u,v;\hat{\beta},\psi) - \iint_{\|u-v\| \leq R} \rho^{(2)}(u,v;\hat{\beta},\psi)\,\mathrm{d}u\,\mathrm{d}v$$

NB: translation invariance for pair correlation not required.

Two-step estimation

Obtain estimates $(\hat{\beta}, \hat{\psi})$ in two steps

1. obtain $\hat{\beta}$ using composite likelihood
2. obtain $\hat{\psi}$ using minimum contrast/second order composite likelihood

Outline:

First-order estimating equations

Score of first order composite likelihood is

$$\frac{\mathrm{d}}{\mathrm{d}\beta} \log CL(\beta) = \sum_{u \in \mathbf{X} \cap W} \frac{\rho'_\beta(u)}{\rho_\beta(u)} - \int_W \rho'_\beta(u) \mathrm{d}u$$

Special case of unbiased *first-order* estimating function

$$u_f(\beta) = \sum_{u \in \mathbf{X} \cap W} f_\beta(u) - \int_W f_\beta(u) \rho_\beta(u) \mathrm{d}u$$

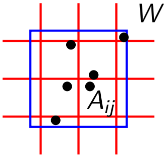
with

$$f_\beta(u) = \frac{\rho'_\beta(u)}{\rho_\beta(u)}$$

This is optimal choice for Poisson process (MLE) but what is optimal in the clustered case ?

Asymptotic results - first order estimating function

Divide $\mathbb{R}^2$ into quadratic cells
$A_{ij} = [i, i+1[\times [j, j+1[$



Then

$$u_f(\beta) = \sum_{ij : A_{ij} \subseteq W} U_{ij}$$

where

$$U_{ij} = \sum_{u \in \mathbf{X} \cap A_{ij}} f_\beta(u) - \int_{A_{ij}} f_\beta(u) \rho_\beta(u) \mathrm{d}u$$

Assuming $\mathbf{X}$ is mixing, $\{U_{ij}\}_{ij}$ mixing random field and

$$|W|^{-1/2} u_f(\beta) \approx N(0, \Sigma_f)$$

(CLT for mixing random field $\{U_{ij}\}_{ij}$).

Mixing

Consider $E_1, E_2 \subseteq \mathbb{R}^2$ and point configurations F_1 and F_2.

Need polynomial decay of

$$|P(\mathbf{X} \cap E_1 \in F_1, \mathbf{X} \cap E_2 \in F_2) - P(\mathbf{X} \cap E_1 \in F_1)P(\mathbf{X} \cap E_2 \in F_2)|$$

as function of distance between E_1 and E_2.

This can easily be verified for a shot-noise process where the kernel density decays fast enough.

Asymptotic results cntd.

Estimate $\hat{\beta}$ solves

$$u_f(\beta) = 0$$

And (Taylor)

$$u_f(\beta) \approx |W|(\hat{\beta} - \beta)S_f$$

where

$$S_f = -\mathbb{E}\frac{\mathrm{d}}{\mathrm{d}\beta^\mathsf{T}} u_f(\beta)/|W|$$

It follows that

$$\hat{\beta} \approx N(\beta, V_f/|W|)$$

where

$$V_f = S_f^{-1}\Sigma_f S_f^{-1}$$

Optimal first-order estimating equation

Optimal choice of f_β: smallest asymptotic variance

$$V_f = S_f^{-1}\Sigma_f S_f^{-1}$$

.

Optimal choice of f_β is solution of Fredholm equation

$$f_\beta(u) + \int_W t(u,v) f_\beta(v)\mathrm{d}u = \frac{\rho'_\beta(u)}{\rho_\beta(u)}, \quad u \in W,$$

where integral equation kernel is

$$t(u,v) = \rho(v)[g(u-v) - 1]$$

Numerical solution of Fredholm equation

Divide W into cells C_i of area $w_i = |C_i|$ and with representative points $u_i \in B_i$.

Matrix-vector approximation of Fredholm equation

$$f_\beta(u) + \int_W t(u,v) f_\beta(v) \mathrm{d}u = \frac{\rho'_\beta(u)}{\rho_\beta(u)}$$

becomes

$$\bar{f} + T\bar{f} = R^{-1}D \Leftrightarrow \bar{f} = (I+T)^{-1}R^{-1}D$$

where

$$\bar{f} = [f_\beta(u_i)]_i \quad T = \left[w_j \rho_\beta(u_j)[g(u_i,u_j)-1]\right]_{ij}$$

and

$$R = \mathrm{diag}\left[w_i \rho_\beta(u_i)\right] \quad D = \left[w_i \mathrm{d}\rho(u_i)/\mathrm{d}\beta_l\right]_{il}$$

Let N_i count of points in C_i

$$\mu_i = \mathbb{E}N_i = w_i \rho(u_i),$$

and

$$V = R(I + T) = [V_{ij}]_{ij}$$

where

$$V_{ii} = \mathbb{V}\mathrm{ar} N_i = \mu_i + \mu_i^2[g(u_i, u_i) - 1]$$

$$V_{ij} = \mathbb{C}\mathrm{ov}[N_i, N_j] = \mu_i \mu_j [g(u_i, u_j) - 1]$$

Then

$$\bar{f} = (I + T)^{-1} R^{-1} D = V^{-1} D$$

where

$$D = \left[w_i \mathrm{d}\rho_\beta(u_i)/\mathrm{d}\beta_l \right]_{il} = \left[\mathrm{d}\mu(u_i)/\mathrm{d}\beta_l \right]_{il}$$

Inserting stepfunction approximation $\bar{f}$ into

$$\sum_{u \in \mathbf{X}} f_\beta(u) - \int_W f_\beta(u) \rho_\beta(u) \mathrm{d}u$$

yields

$$\begin{aligned}\sum_i \sum_{u \in \mathbf{X} \cap B_i} \bar{f}_i - \sum_i \bar{f}_i w_i \rho_\beta(u_i) =& \sum_i (N_i - \mu_i) \bar{f}_i \\ =& (N - \mu) \bar{f} \\ =& (N - \mu) V^{-1} D\end{aligned}$$

where

$$N = [N_1, \ldots, N_m]$$

$(N - \mu) V^{-1} D$ is the *quasi-likelihood* (generalized estimating equation) based on count data vector N.

Practical implementation: IGLS

In practice: numerically approximated Fredholm equation $\Rightarrow$ quasi-likelihood for N.

Pair correlation function inside V estimated by e.g. minimum contrast.

Solve for β using iterative generalized least squares:

$$(\beta^{(l+1)} - \beta^{(l)}) D(\beta^{(l)})^{\mathsf{T}} V(\beta^{(l)})^{-1} D(\beta^{(l)}) = (N - \mu(\beta^{(l)})) V(\beta^{(l)})^{-1} D(\beta^{(l)})$$

One issue: use fine discretization (large m) $\Rightarrow$ V highdimensional matrix - e.g. V 10000×10000.

Use tapering and sparse matrix Cholesky from `Matrix` library in R.

Simulation study

Consider variance of $\hat{\beta}$ obtained from either composite likelihood or GEE.

Reduction in variance for GEE relative to composite likelihood: 6% to 65%.

Large reductions when strong clustering and strong inhomogeneity.

Outline:

Background in tropical forest ecology

Intensity function and second order summary statistics

Poisson and Cox processes

Estimation

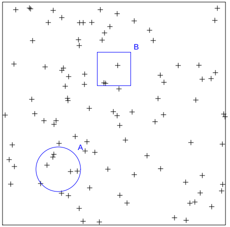
Optimal first-order estimating equations - quasi-likelihood

Decomposition of variance

Decomposition of variance for a count

Prediction of count $N(B)$ given Λ :

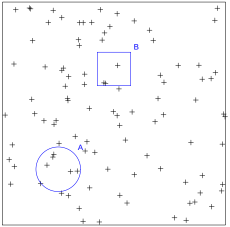
$$\hat{N}(B) = \mathbb{E}[N(B)|\Lambda] = \int_B \Lambda(u)\mathrm{d}u$$



Decomposition of variance for a count

Prediction of count $N(B)$ given Λ :

$$\hat{N}(B) = \mathbb{E}[N(B)|\Lambda] = \int_B \Lambda(u)\mathrm{d}u$$



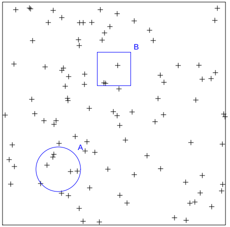
$\mathbb{V}\mathrm{ar} N(B) =$

variation $=$

Decomposition of variance for a count

Prediction of count $N(B)$ given Λ :

$$\hat{N}(B) = \mathbb{E}[N(B)|\Lambda] = \int_B \Lambda(u)\mathrm{d}u$$



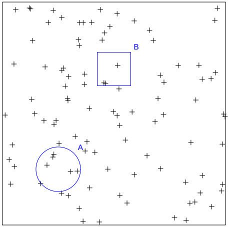
$$\mathbb{V}\mathrm{ar} N(B) = \mathbb{V}\mathrm{ar} \hat{N}(B)$$

variation = variation Λ

Decomposition of variance for a count

Prediction of count $N(B)$ given Λ :

$$\hat{N}(B) = \mathbb{E}[N(B)|\Lambda] = \int_B \Lambda(u)\mathrm{d}u$$



$$\mathbb{V}\mathrm{ar} N(B) = \mathbb{V}\mathrm{ar} \hat{N}(B) + \mathbb{V}\mathrm{ar}\left[N(B) - \hat{N}(B)\right]$$

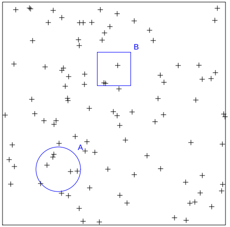
variation = variation Λ + 'Poisson noise'

Further, $\hat{\Lambda}(u) = \mathbb{E}[\Lambda(u)|Z]$

Decomposition of variance for a count

Prediction of count $N(B)$ given Λ :

$$\hat{N}(B) = \mathbb{E}[N(B)|\Lambda] = \int_B \Lambda(u)\mathrm{d}u$$



$$\mathbb{V}\mathrm{ar} N(B) = \mathbb{V}\mathrm{ar} \hat{N}(B) + \mathbb{V}\mathrm{ar}\left[N(B) - \hat{N}(B)\right]$$

variation = variation Λ + 'Poisson noise'

Further, $\hat{\Lambda}(u) = \mathbb{E}[\Lambda(u)|Z]$

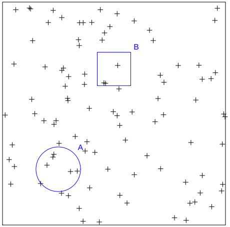
$\mathbb{V}\mathrm{ar}\Lambda(u) =$

structured variation =

SST =

Decomposition of variance for a count

Prediction of count $N(B)$ given Λ :



$$\hat{N}(B) = \mathbb{E}[N(B)|\Lambda] = \int_B \Lambda(u)\mathrm{d}u$$

$$\mathbb{V}\mathrm{ar}N(B) = \mathbb{V}\mathrm{ar}\hat{N}(B) + \mathbb{V}\mathrm{ar}\left[N(B) - \hat{N}(B)\right]$$

variation $=$ variation $\Lambda +$ 'Poisson noise'

Further, $\hat{\Lambda}(u) = \mathbb{E}[\Lambda(u)|Z]$

$$\mathbb{V}\mathrm{ar}\Lambda(u) = \mathbb{V}\mathrm{ar}\hat{\Lambda}(u)$$

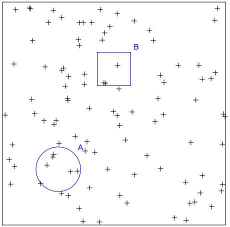
structured variation $=$ variation due to environment

SST $=$ SSR

Decomposition of variance for a count

Prediction of count $N(B)$ given Λ :

$$\hat{N}(B) = \mathbb{E}[N(B)|\Lambda] = \int_B \Lambda(u)\mathrm{d}u$$



$$\mathbb{V}\mathrm{ar} N(B) = \mathbb{V}\mathrm{ar} \hat{N}(B) + \mathbb{V}\mathrm{ar}\left[N(B) - \hat{N}(B)\right]$$

variation $=$ variation Λ $+$ 'Poisson noise'

Further, $\hat{\Lambda}(u) = \mathbb{E}[\Lambda(u)|Z]$

$$\mathbb{V}\mathrm{ar}\Lambda(u) = \mathbb{V}\mathrm{ar}\hat{\Lambda}(u) + \mathbb{V}\mathrm{ar}\left[\Lambda(u) - \hat{\Lambda}(u)\right]$$

structured variation $=$ variation due to environment $+$ other sources

SST $=$ SSR $+$ SSE

Measure of influence of environmental covariates Z:

$$R^2 = \frac{SSR}{SST} = \frac{\mathbb{V}\mathrm{ar}\mathbb{E}[\Lambda(u)|Z]}{\mathbb{V}\mathrm{ar}\Lambda(u)}$$

(right hand side does not depend on u in case of stationary environment)

Additive and log-linear models

Model influence of environment using linear model:

$$\tilde{Z}(u) = \beta Z(u)^{\mathsf{T}}$$

Additive model:

$$\Lambda(u) = \beta Z(u)^{\mathsf{T}} + \Lambda_0 = \tilde{Z}(u) + \Lambda_0(u)$$

(superposition of point processes with intensity functions $\tilde{Z}$ and Λ_0
- convenient for variance decomposition)

Additive and log-linear models

Model influence of environment using linear model:

$$\tilde{Z}(u) = \beta Z(u)^{\mathsf{T}}$$

Additive model:

$$\Lambda(u) = \beta Z(u)^{\mathsf{T}} + \Lambda_0 = \tilde{Z}(u) + \Lambda_0(u)$$

(superposition of point processes with intensity functions $\tilde{Z}$ and Λ_0 - convenient for variance decomposition)

Log-linear model

$$\Lambda(u) = \exp[\beta Z(u)^{\mathsf{T}}]\Lambda_0(u) = \exp[\tilde{Z}(u)]\Lambda_0(u)$$

(independent thinning of point process $\mathbf{X}_0$ with intensity function Λ_0)

R^2 for additive and log-linear models

Additive $\Lambda(u) = \tilde{Z}(u) + \Lambda_0(u)$:

$$R^2 = \frac{\sigma^2_{\tilde{Z}}}{\sigma^2_{\tilde{Z}} + \sigma^2_0}$$

$$\sigma^2_{\tilde{Z}} = \mathbb{V}\mathrm{ar}\tilde{Z}(u) \quad \text{and } \sigma^2_0 = \mathbb{V}\mathrm{ar}\Lambda_0(u)$$

R^2 for additive and log-linear models

Additive $\Lambda(u) = \tilde{Z}(u) + \Lambda_0(u)$:

$$R^2 = \frac{\sigma^2_{\tilde{Z}}}{\sigma^2_{\tilde{Z}} + \sigma^2_0}$$

$$\sigma^2_{\tilde{Z}} = \mathbb{V}\mathrm{ar}\tilde{Z}(u) \quad \text{and } \sigma^2_0 = \mathbb{V}\mathrm{ar}\Lambda_0(u)$$

Log-linear $\Lambda(u) = \exp[\tilde{Z}(u)]\Lambda_0(u)$:

$$R^2 = \frac{\sigma^2_{\exp \tilde{Z}}}{\sigma^2_{\exp \tilde{Z}} + \sigma^2_0\left[\sigma^2_{\exp \tilde{Z}} + \mu^2_{\exp \tilde{Z}}\right]}$$

$$\sigma^2_{\exp \tilde{Z}} = \mathbb{V}\mathrm{ar}\exp[\tilde{Z}(u)] \quad \text{and } \mu_{\exp \tilde{Z}} = \mathbb{E}\exp[\tilde{Z}(u)]$$

Estimation: environmental variances

Z observed on grid $G = \{u_i\}_{i=1,\ldots,M}$

Simple empirical estimates, e.g.

$$\widehat{\sigma}^2_{\widetilde{Z}} = \frac{1}{M} \sum_{u \in G} \left(\widehat{\widetilde{Z}}(u) - \widehat{\mu}_{\widetilde{Z}} \right)^2$$

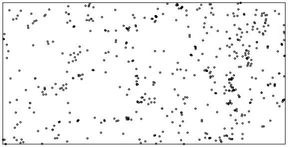
where

$$\widehat{\mu}_{\widetilde{Z}} = \frac{1}{M} \sum_{u \in G} \widehat{\widetilde{Z}}(u)$$

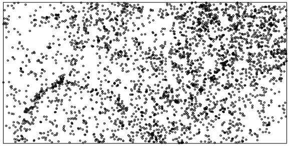
and $\widehat{\widetilde{Z}}(u) = \widehat{\beta} Z(u)^{\mathsf{T}}$.

Three species with different modes of seed dispersal:

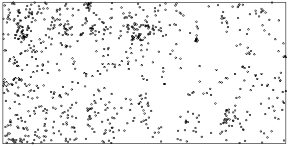
Acalypha Diversifolia explosive capsules



Capparis Frondosa bird/mammal



Loncocharpus Heptaphyllus wind



Estimation using first and second-order composite likelihood.

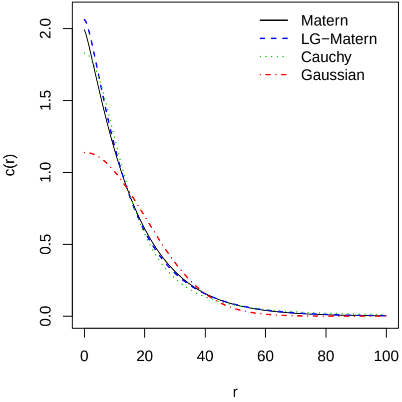
(additive model not second-order reweighted stationary so minimum contrast estimation does not work)

Results for rain forest data

Species	model for Λ	c_0	$\Delta CL_2(\widehat{\psi}\vert\widehat{\beta})$	R^2
Acalypha	log-linear	'Gaussian'	27438	0.01
		Matérn	28507	0.01
	additive	'Gaussian'	0	0.01
		Matérn	1129	0.02
Loncocharpus	log-linear	'Gaussian'	82007	0.11
		Matérn	82327	0.06
	additive	'Gaussian'	0	0.17
		Matérn	702	0.09
Capparis	log-linear	'Gaussian'	5013	0.28
		Matérn	5343	0.11
	additive	'Gaussian'	0	0.29
		Matérn	466	0.12

Some conclusions

Covariance functions for loncocharpus



Best fit with Matérn (heavy tails for covariance/cluster density).

Best fit with log-linear model (interpretation in terms of survival).

Largest R^2 for Capparis (bird/mammal seed dispersal), smallest for Acalypha (explosive capsules).

Thanks to co-authors Yongtao Guan and Abdollah Jalilian (optimal first-order estimating equations and decomposition of variance)

Thanks to co-authors Yongtao Guan and Abdollah Jalilian (optimal first-order estimating equations and decomposition of variance)

and

Thanks for your attention !