LØSNINGSFORSLAG

Emne SIF5064 ROMLIG STATISTIKK

Tirsdag 14. mai 2002

Henning Omre

Oppgave 1 Kontinuerlige Felt

a) $C(\tau)$ må være en positiv definitt funksjon dvs:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j C(|x_i - x_j|) \geq 0 \quad ; \text{ alle } n \geq 1 \quad ; \text{ alle } \alpha = (\alpha_1, .., \alpha_n) \quad ; \text{ alle config } (x_1, .., x_n)$$

$$\mathrm{Cov}\{R(x'), R(x'')\} = E\{(R(x') - \mu(x'))(R(x'') - \mu(x''))\}$$
$$= E\{R(x')R(x'')\} - \left(\sum_i a_i g_i(x')\right)\left(\sum_j a_j g_j(x'')\right)$$

Problemet er at $a = (a_0, \ldots, a_3)$ er ukjente!

Estimér $\hat{a} = (\hat{a}_0, .., \hat{a}_3)$ vha regresjon $\sum_i a_i g_i(x)$ på R°.

For en gitt avstand Δx har en da estimatoren:

$$\hat{C}(\Delta x) = \frac{1}{N_{\Delta x}} \sum_{ij \in A_{\Delta x}} \left[R(x_i) R(x_j) - \left(\sum_k \hat{a}_k g_k(x_i)\right)\left(\sum_\ell \hat{a}_\ell g_\ell(x_j)\right)\right]$$

hvor

$A_{\Delta x} : \{ i,j \mid \Delta x + \delta < x_i - x_j < \Delta x - \delta \}$

$N_{\Delta x} \doteq \# A_{\Delta x}$

En kan bestemme $\hat{C}(\Delta x)$ for fler Δx, plotte dem, og tilpasse en pos. def. funksjon $\hat{C}(\tau); \tau \geq 0$.
Denne estimatoren er selvsagt ikke forventningsrett.
Alternativt kan en anta at $\{R(x); x \in D\}$ er et Gaussisk RF og bruke en maximum likelihood estimator mhp R°.

c) Definér prediktoren

$$\hat{R}(x_0) = \sum_j^{50} \alpha_j R(x_j)$$

Forventningsretthet:

$$E\{\hat{R}(x_0) - R(x_0)\} = 0 \Rightarrow \sum_j^{50} \alpha_j g_l(x_j) = g_l(x_0) \; ; \; l = 0, \dots, 3$$

Prediksjonsvarians:

$$Var\{\hat{R}(x_0) - R(x_0)\} = C(0) - 2\sum_j \alpha_j C(x_0 - x_j) + \sum_j \sum_k \alpha_j \alpha_k C(x_j - x_k)$$

Minimeringssystem som må løses:

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}} \left[C(0) - 2\sum_j \alpha_j C(x_0 - x_j) + \sum_j \sum_k \alpha_j \alpha_k C(x_j - x_k) \right]$$

bibetingelse:

$$\sum_j \alpha_j g_l(x_j) = g_l(x_0) \; ; \; l = 0, \dots, 3$$

d) Med R^{oo} får en forventningsretthetskravet i b):

$$\left.\begin{aligned} \alpha_1 g_0(x_1) + \alpha_2 g_0(x_2) &= g_0(x_0) \\ &\vdots \\ \alpha_1 g_3(x_1) + \alpha_2 g_3(x_2) &= g_3(x_0) \end{aligned}\right\}$$

Fire krav og to variable. Ikke mulig å sikre forventningsretthet!!

Redefiner prediktoren:

$$\hat{R}(x_0) = \sum_{i=0}^{3} \mu_{A_i} g_i(x_0) + \sum_{j=1}^{2} \alpha_j \left[R(x_j) - \sum_{i=0}^{3} \mu_{A_i} g_i(x_j) \right]$$

Herav

$$E\{\hat{R}(x_0) - R(x_0)\} = 0$$ – automatisk forventningsrett

Dette oppstår fordi $\sum a_i g_i(x)$ kan 'sentreres' om $\sum \mu_{A_i} g_i(x)$ fordi $E\{A\} = \mu_A$ er kjent!!

Prediksjonsvariansen:

$$\mathrm{Var}\{\hat{R}(x_0) - R(x_0)\}$$

$$= G(0) + g(x_0)^T \Sigma_A g(x_0) - 2\sum_j \alpha_j \left[G(x_0 - x_j) + g(x_j)^T \Sigma_A g(x_0) \right]$$

$$+ \sum_j \sum_k \alpha_j \alpha_k \left[G(x_j - x_k) + g(x_j)^T \Sigma_A g(x_k) \right]$$

Minimeringsproblemet er definert som:

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}} \ \mathrm{Var}\{\hat{R}(x_0) - R(x_0)\}$$

altså uten bibetingelser siden forventningsretthet er sikret ved konstruksjon av $\hat{R}(x_0)$. Alle variable i minimeringen er antatt kjent.

Oppgave 2 Hendelses Felt

a) En estimator er:

$$\hat{\lambda} = \frac{1}{7}\sum_{i=1}^{7} \hat{\lambda}_i = \frac{1}{7}\sum_{i=1}^{7} \frac{N_i}{\pi r_i^2} \rightarrow \hat{\lambda} = 0.18\,[\text{pkt}/\text{cm}^2]$$

en annen:

$$\hat{\hat{\lambda}} = \frac{\sum_i N_i}{\pi \sum_i r_i^2} \rightarrow \hat{\hat{\lambda}} =$$

b) Definer L-funksjonen:

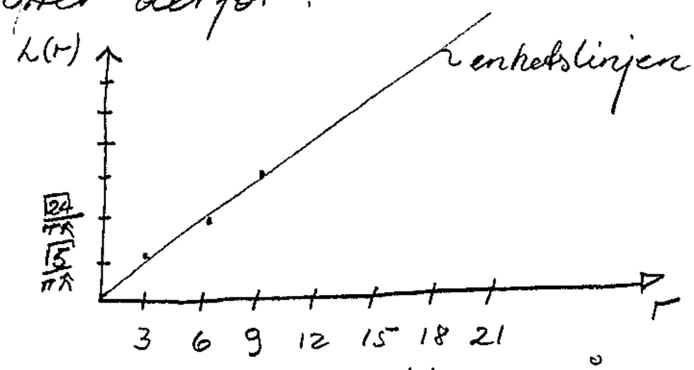
$$L(r) = \left[E\{N(r B_0)\} / \pi\lambda \right]^{1/2}$$

dvs estimator:

$$\hat{L}(r_i) = \left[\frac{N_i}{\pi \hat{\lambda}}\right]^{1/2} \quad ; \quad i = 1, \ldots, 7$$

Merk! For Poisson RF har en $L(r) = r$.

Plotter derfor:



Dersom estimatene ligger på enhetslinjen indikerer det at det er et Poisson RF.

For å vurdere signifikans kan en:

- simulere Poisson RF med intensitet $\hat{\lambda}$.
- simulere sampling prosedyren
- plotte simulerte observasjoner

Dette gjøres mange ganger – variabiliteten sammenliknes med avviket fra enhetslinjen som observasjonene gav.

g) Husk

$$\mathrm{Prob}\{N(r\mathcal{B}_0)=i\}=\frac{(\lambda\pi r^2)^i}{i!}e^{-\lambda\pi r^2} \quad ; i=0,1,\ldots$$

Likelihood for samplet:

$$L(\lambda)=\mathrm{Prob}\{N(3\mathcal{B}_0)=5 \cap \ldots \cap N(21\mathcal{B}_0)=248 ; \lambda\}$$

$$=\lambda^{(5+\cdots+248)}\cdot\frac{(\pi 3^2)^5\ldots(\pi 21^2)^{248}}{5!\quad 248!}\cdot e^{-\lambda\pi(3^2+\cdots+21^2)}$$

Maximum likelihood estimatoren er:

$$\tilde{\lambda}=\underset{\lambda}{\mathrm{argmax}}\; L(\lambda)=\underset{\lambda}{\mathrm{argmax}}\;\ln L(\lambda)$$

$$=\underset{\lambda}{\mathrm{argmax}}\{\underbrace{c_1}_{(5+\cdots+248)}\ln\lambda+\ln c_2-\lambda\pi\underbrace{c_3}_{(3^2+\cdots+21^2)}\}$$

$$\Downarrow$$

$$\frac{d}{d\lambda}(c_1\ln\lambda+\ln c_2-\lambda\pi c_3)=0$$

$$\Downarrow$$

$$\tilde{\lambda}=\frac{c_1}{\lambda c_3}\longrightarrow \text{tilsvarer estimator } \hat{\hat{\lambda}} \text{ i a).}$$

Alternativ estimator under Poisson RF antakelse:

$$L(r)=\frac{1}{\lambda^{1/2}}\underbrace{\left[E\{N(r\mathcal{B}_0)\}/\pi\right]^{1/2}}_{L_\lambda(r)}=r$$

$$\Downarrow$$

$$L_\lambda(r)=\lambda^{1/2}r$$

Merk at $\hat{L}_\lambda(r_i)$; $i=1,\ldots,7$ kan beregnes fra tabellen.

Tilpass vha lineær regresjon ar til $\hat{L}_\lambda(r_i)$; $i=1,\ldots,7$ og bestem $\hat{a}$.

Da har en

$$\lambda^*=\hat{a}^2.$$

Oppgave 3 Mosaikk Felt

a) Likelihood funksjonen er:

$$L(\beta) = \text{Prob}\{L = l^\circ ; \beta\} = c(\beta) \cdot \exp\Big\{\beta \sum_{x \sim y} I(l_x^\circ = l_y^\circ)\Big\}$$

med

$$c(\beta) = \Big[\sum \exp\Big\{\beta \sum_{x\sim y} I(l_x = l_y)\Big\}\Big]^{-1}$$

alle mulige konfigurasjoner av $L = l$!

Hvis griddet er n_D, vil konfigurasjonssummen være over 2^{n_D} ledd – stort tall. I tillegg må en summere over alle naboer i denne summen igjen! $c(\beta)$ kan altså vanskelig bestemmes eller estimeres!

Pseudo-likelihood funksjonen:

$$L_p(\beta) = \prod_{x \in \mathcal{L}_D} \text{Prob}\{L_x = l_x^\circ \mid L_y = l_y^\circ ; y \in N_x ; \beta\}$$

naboskap til x

$$= \prod_{x \in \mathcal{L}_D} c_x(\beta) \exp\Big\{\beta \sum_{y \in N_x} I(l_x^\circ = l_y^\circ)\Big\}$$

med

$$c_x(\beta) = \Big[\exp\Big\{\beta \sum_{y\in N_x} I(-1 = l_y^\circ)\Big\} + \exp\Big\{\beta \sum_{y \in N_x} I(1 = l_y^\circ)\Big\}\Big]^{-1}$$

grei å regne ut for hver $x \in \mathcal{L}_D$ for gitt β.

Maximum pseudo-likelihood estimatoren blir da:

$$\hat\beta = \underset{\beta}{\text{argmax}}\ L_p(\beta) = \underset{\beta}{\text{argmax}}\ \ln L_p(\beta)$$

$$= \underset{\beta}{\text{argmax}} \Big\{\sum_x \ln c_x(\beta) + \beta \sum_x \sum_{y \in N_x} I(l_x^\circ = l_y^\circ)\Big\}$$

$$\Downarrow$$

$$-\sum_x \frac{1}{c_x(\hat\beta)} \frac{d\, c_x(\hat\beta)}{d\hat\beta} = \sum_x \sum_{y \in N_x} I(l_x^\circ = l_y^\circ)$$

definerer $\hat\beta$.