

Venables & Dichmont (2004)
LMs, GLMs, GAMs e GLMM
Review paper on Fisheries Research

Seminário de discussão

LEG: Laboratório de Estatística e Geoinformação
Universidade Federal do Paraná

1 Models for fisheries: GAMs, GLMs & GLMMs

Número especial da Fisheries research

aplicações dos modelos e revisão da metodologia

1.1 Contexto do número especial

não linearidade *versus* conveniência computacional:

associação entre abundância e variáveis biológicas e ambientais

não normalidade por assimetria, contagem, etc

LM , GLM , GEE, GAM, GLMM, modelos de espaço de estados e espaciais
cada vez mais usados

Grandes tópicos

- padronização CPUE
- outros tópicos (revisão e aplicações diversas)

LM e extensões continuarão largamente utilizados

conjectura: GLMM e GEE como escolhas do futuro

2 Modelos:

O bom, o mau e o feio?
Da conveniência à melhor descrição da realidade

2.1 A concepção do modelo

resposta e variáveis de estímulo

determinístico *versus* aleatório: como descrever (e interpretar!) variação ou aleatoriedade

Seja $x = (x_1, x_2, \dots, x_k)$ e z variável centrada em zero

Modelo geral: $y = f(x, z)$ como sistema contínuo de variação suave

Aproximação local: Expansão por Taylor em torno de $(x_0, 0)$:

$$y = \beta_0 + \sum_{j=1}^k \beta_j (x_j - x_{j0}) + \sigma z$$

LM!!!

2.2 Mais geral?

considere uma vizinhança maior, expansão de segunda ordem

interpretação: acomoda-se termo quadrático, interações

maior susceptibilidade a desvios de normalidades, assimetria, curtose, etc

modelo realístico *versus* coerente: restrições: origem, bordas, etc

2.3 GLMs

acomodam: heterogeneidade de variância, assimetria, não normalidade

verossimilhança aproximadamente quadrática

em geral boas propriedades para inferência

teoria bem estabelecida, implementações disponíveis, etc

2.4 GAMs

Acomodando curvaturas: polinômios *versus* splines

termos, termos e mais termos ...

training e testing

comentário sobre aleatoriedade:

aleatoriedade não é a mera ausência de padrão. Aleatoriedade pode em geral mostrar padrões bem definidos! O problema é, na próxima vez o padrão será o mesmo, ou pode ser completamente diferente?

GAMs automatizam escolhas em GLMs

Preços a pagar: (*there is no free lunch*)

- não mais a bela e simples inferência baseada em verossimilhança como em GLMs
- penalização, pesos e validação cruzada

2.5 GLMM

Mudança/extensão de GLMs mais conceitual, maior e profunda
diversos pontos de entrada para componentes estocásticos,

$$y = f(x, z_1, z_2, \dots)$$

componentes de variância

natural para grupos

1 parâmetro por grupo ou amostra de uma distribuição (Normal?)

estimação *versus* predição de parâmetros

relações com GAMs: penalização e restrição nos parâmetros

achatamento (*shrinkage*)

emprestando informação (*borrowing strenght*)

2.6 Comentários

resumindo: modelos lineares são portanto empíricos e locais
(e ainda contam com a normalidade)!!

transformação de dados *versus* da média

o problema em retornar à escala original para discussão dos resultados

GLM e resultados assintóticos

extensões de quase-verossimilhança

generalização por splines e a necessidade de informação contextual

GAMs trazem critérios

interações e relações podem ficar obscuras

GLMM: elegância e praticidade