Topics in Statistics : Graphical Models

Spring 2007

Richard Gill

http://www.math.leidenuniv.nl/~gill/teaching/graphical

Preliminaries on conditioning

- Assume all random variables take values in Polish spaces and hence all joint distributions can be disintegrated into marginals and families of conditionals
- $X \perp Y \mid Z$, X indep of Y given Z ,

$\Leftrightarrow$ $\mathrm{Law}(X,Y|Z=z) = \mathrm{Law}(X|Z=z) \otimes \mathrm{Law}(Y|Z=z)$ P_Z-a.e.(z)

$\Leftrightarrow$ $P_{X,Y|Z=z} = P_{X|Z=z} \otimes P_{Y|Z=z}$

Preliminaries on conditioning

- Say P is positive iff $\otimes_i P_i \ll P$ (abs cts/dominated)

  $\Leftrightarrow P(\forall\, i\ X_i \in B_i) = 0 \Rightarrow \Pi_i P(X_i \in B_i) = 0$

  $\Leftrightarrow \Pi_i P(X_i \in B_i) > 0 \Rightarrow P(X_i \in B_i\ \forall\, i) > 0$

  $\Leftrightarrow$ P has a strict +ve density wrt a product meas.

  $\Leftrightarrow$ P has no nontrivial deterministic relations

- We always have $P \ll \otimes_i P_i$, P always has a density with respect to a product measure. So P is positive iff $P \equiv \otimes_i P_i$

Graphoid axioms

ternary relation R on subsets of a finite set V

"C blocks B-to-A" , $A \leftarrow B \mid C$ or $A \overset{C}{\not\leftarrow} B$

- (G1) $A \leftarrow B \mid C \Rightarrow B \leftarrow A \mid C$
- (G2) $A \leftarrow B \mid C$, $U \subseteq A \Rightarrow U \leftarrow B \mid C$
- (G3) $A \leftarrow B \mid C$, $U \subseteq B \Rightarrow A \leftarrow B \mid C \cup U$
- (G4) $A \leftarrow C \mid B$, $A \leftarrow D \mid B \cup C \Rightarrow A \leftarrow C \cup D \mid B$
- (G5) A, B, C disjoint , $A \leftarrow B \mid C$, $A \leftarrow C \mid B$
  $\Rightarrow A \leftarrow B \cup C \mid \emptyset$

Graphoids, graph sep. in undir. graphs, conditional independence

- Th. 1 : a 3^{ary} relation R satisfies (G1)-(G5)
  $\Leftrightarrow$ $\exists$ graph G=(V,E) s.t. R is G-separation
- Th. 2 : Given P define R: $A \leftarrow B|C \Leftrightarrow X_A \perp X_B | X_C$
  then R satisfies (G1)-(G4), and (G5) if P is +ve
- Th. 3 : Given G $\exists$ +ve P s.t. S G-separates
  A from B $\Leftrightarrow$ $X_A \perp X_B | X_S$

Impossibility of finite axiomatisation (Studenyi)

- Suppose (C1)–(Cn) are nontrivial true implications for $R=(\cdot\perp\cdot|\cdot)$

  Then $\exists$ nontrivial true implication (D) for R

  such that (C1)–(Cn) $\not\Rightarrow$ (D)

- However as we saw, all +ve P satisfy all (D) implied by (G1)–(G5)

Markov Properties wrt undirected graphs

$G=(V,E)$ $X = (X_\alpha : \alpha \in V)$ Write α for X_α etc

Consider

- (P) $\alpha \not\sim \beta \Rightarrow \alpha \perp \beta \mid V \setminus \{\alpha,\beta\}$
- (L) $\alpha \perp V \setminus cl(\alpha) \mid ne(\alpha)$
- (G) $A \overset{S}{\not\leftrightarrow} B \Rightarrow A \perp B \mid S$
- (F) $f(x) = \Pi_{A \text{ complete}} \Phi_A(x_A)$
  (f density of P wrt some product measure)

Markov Properties wrt undir graphs

- Thm 1: (F)actor ⇒ (G)lobal ⇒ (L)ocal ⇒ (P)airwise Markov property
- Thm 2: If f positive, the (P)airwise => (F)actor

(Hammersley-Clifford theorem)

Proofs

- Theorem 1. Trivial using (G1)-(G4) together with density factorization criterion of conditional independence
[wlog assume A, B, S a partition of V; note that if C is complete and S separates A from B then C can't hit both A and B]

Proofs (continued)

- Theorem 2. Lemma : Möbius inversion :
  $H_A = \Sigma_{B \subseteq A} G_B \quad \Leftrightarrow \quad G_B = \Sigma_{A \subseteq B} (-1)^{|B \setminus A|} H_A$

- Now pick x^* and define $H_A(x_A) = \log(f(x_A, x_{V \setminus A}^*))$ so that $G_B = \Sigma_{A \subseteq B} (-1)^{|B \setminus A|} H_A$ satisfies $H_A = \Sigma_{B \subseteq A} G_B$,

  in particular $\log(f(x)) = \Sigma_B \, G_B(x_B)$

- Suppose B not complete , $\alpha \nsim \beta \in B$. Then $G_B(x_B) =$

$$\Sigma_{A \subseteq B \setminus \{\alpha,\beta\}} (-1)^{|B \setminus A|} \log \frac{f(x_\alpha^*, x_\beta^*, x_A, x_R^*) / f(x_\alpha, x_\beta^*, x_A, x_R^*)}{f(x_\alpha^*, x_\beta, x_A, x_R^*) / f(x_\alpha, x_\beta, x_A, x_R^*)}$$

(continued next slide)

Proofs (continued)

(f +ve)

- But $A \perp B \mid C \iff \dfrac{f(x_A^*,x_B^*,x_C)/f(x_A,x_B^*,x_C)}{f(x_A^*,x_B,x_C)/f(x_A,x_B,x_C)} = 1$

  for all ...