Möbius inversion

Let Ψ be a real valued function defined on the subsets of a finte set $\mathbb{I}$. For each $A \subseteq \mathbb{I}$ define

$$\Phi(A) = \sum_{B \subseteq A} \Psi(B),$$

the sum running over all subsets (including the empty set) of A. Then

<1> $$\Psi(A) = \sum_{B \subseteq A} (-1)^{\#(A \setminus B)} \Phi(B),$$

where $\#E$ for a cardinality of a set E.

Proof. First note that $\sum_{B \subseteq A} (-1)^{\#(A \setminus B)} = 0$ for each nonempty set A, because 2^{-n} times the left-hand side equals $\mathbb{P}(-1)^{\sum_{i \in A} Z_i} = \prod_i \mathbb{P}(-1)^{Z_i}$, where the Z_i are independent random variables taking the values 0 and 1 each with probability $1/2$. If A were empty, the sum would equal 1.

The right-hand side of <1> equals

$$\sum_{B \subseteq A} \sum_{C \subseteq B} (-1)^{\#(A \setminus B)} \Psi(C) = \sum_{C \subseteq A} \Psi(C) \sum_{E \subseteq A \setminus C} (-1)^{\#E}$$

The inner sum equals zero except when $A = C$. □

Markov random fields and Gibbs distributions

Let $\mathbb{Q}$ be a probability measure defined on a product space $\mathcal{X} = \otimes_{i \in \mathbb{I}} \mathcal{X}_i$, with $\mathbb{I} = \{1, 2, \ldots, n\}$ a finite index set. Suppose $\mathbb{I}$ is the set of nodes of a graph. Write X_i for the ith coordinate map. Assume $\mathbb{Q}\{\mathbf{x}\} > 0$ for every $\mathbf{x}$ in $\mathcal{X}$. Call $\mathbb{Q}$ a ***Markov random field*** if, for every i, the conditional probability

$$\mathbb{Q}\{X_i = x_i \mid X_j = y_j \text{ for } j \neq i\}$$

depends only on those y_j for which j is a neighbor of i in the graph.

Arbitrarily choose points $z_i \in \mathcal{X}_i$ for $i = 1, 2, \ldots, n$. For each $S \subseteq \mathbb{I}$ and $\mathbf{x} \in \mathcal{X}$ define

$$f_S(\mathbf{x}) = \log \mathbb{Q}\{\mathbf{y}\} \qquad \text{where } y_i = \begin{cases} x_i & \text{if } i \in S \\ z_i & \text{if } i \in S^c \end{cases}$$

Abbreviate $f_{\mathbb{I}}$ to f. For a fixed $\mathbf{x}$, we can think of f_S as an integral of f with respect to a product measure,

$$f_S(\mathbf{x}) = \otimes_i \left(\delta(x_i)\{i \in S\} + \delta(z_i)\{i \in S^c\} \right) f$$

with $\delta(y_i)$ denoting a point mass at y_i, with $y_i = x_i$ or $y_i = z_i$. Rewrite the ith measure in the product as $\delta(z_i) + (\delta(x_i) - \delta(z_i))\{i \in S\}$, then expand the product as

$$\sum_{A \subseteq S} \otimes_i \left(\delta(z_i)\{i \in A^c\} + (\delta(x_i) - \delta(z_i))\{i \in A\} \right) f$$

The sum runs over all subsets of S, including the empty subset. Write $\Psi_A(\mathbf{x})$ for the Ath summand. Thus $f_S(\mathbf{x}) = \sum_{A \subseteq S} \Psi_A(\mathbf{x})$. In particular,

<2> $$\mathbb{Q}\{\mathbf{x}\} = \exp\left(\sum_{A \subseteq \mathbb{I}} \Psi_A(\mathbf{x}) \right)$$

Notice that $\Psi_A(\mathbf{x})$ depends on $\mathbf{x}$ only through the values $\{x_i : i \in A\}$. As a function of $\mathbf{x}$, each Ψ_A has the property

<3> $$\Psi_A(\mathbf{x}) = 0 \qquad \text{if } x_i = z_i \text{ for some } i \text{ in } A$$

<4>**Theorem** *For a Markov random field, the term Ψ_A in <2> is identically zero if A is not a clique of the graph.*

Proof. From the Möbius inversion formula,

<5>
$$\Psi_A(\mathbf{x}) = \sum_{B\subseteq A} (-1)^{\#(A\backslash B)} f_B(\mathbf{x})$$

I will show that there is a cancellation of terms on the right-hand side if A is not a clique.

For simplicity, suppose $1, 2 \in A$ but nodes 1 and 2 are not connected by an edge of the graph, that is, they are not neighbors. Consider the contributions to <5> in pairs $B, \tilde{B}$, where $1 \notin B$ and $\tilde{B} = B \cup \{1\}$. The cardinalities of $A\backslash B$ and $A\backslash \tilde{B}$ differ by 1; the pair $B, \tilde{B}$ contributes $\pm \left(f_{\tilde{B}}(\mathbf{x}) - f_B(\mathbf{x})\right)$ to the sum. Define

$$y_i = \begin{cases} x_i & \text{if } i \in B \\ z_i & \text{if } i \in B^c \end{cases}$$

Then

$$\begin{aligned} f_{\tilde{B}}(\mathbf{x}) - f_B(\mathbf{x}) &= \log \frac{\mathbb{Q}\{X_1 = x_1, X_2 = y_2, \ldots, X_n = y_n\}}{\mathbb{Q}\{X_1 = z_1, X_2 = y_2, \ldots, X_n = y_n\}} \\ &= \log \frac{\mathbb{Q}\{X_1 = x_1 \mid X_2 = y_2, \ldots, X_n = y_n\}}{\mathbb{Q}\{X_1 = z_1 \mid X_2 = y_2, \ldots, X_n = y_n\}} \end{aligned}$$

A common factor of $\mathbb{Q}\{X_2 = y_2, \ldots, X_n = y_n\}$ cancels from numerator and denominator. The Markov property ensures that the conditional probabilities do not depend on the value y_2. The right-hand side of <5> is unchanged if we replace y_2 by z_2. That is, $\Psi_A(\mathbf{x})$ is unchanged if we put x_2 equal to z_2. From <3>, deduce that $\Psi_A(\mathbf{x}) = 0$ for all $\mathbf{x}$, as asserted. □

Remarks:

(a) The converse is easy.

(b) Replace $\delta(x_i)$ by the empirical measure, and $\delta(z_i)$ by the underlying distribution, to get the Hoeffding decomposition for a U-statistic with kernel f.

(c) Is it possible to use Möbius inversion to get simpler proof of the VC Lemma? The pairing in the proof of Theorem <4> reminds me of one proof of the Lemma.

References

Griffeath, D. (1976), *Introduction to Markov Random Fields*, Springer. Chapter 12 of *Denumerable Markov Chains* by Kemeny, Knapp, and Snell (2nd edition).