# COHERENCE AND EXTENSIONS OF STOCHASTIC MATRICES

#### ANGELO GILIO - FULVIO SPEZZAFERRI

In this paper a review of some general results on coherence of conditional probability assessments is given. Then, a necessary and sufficient condition for coherence of two finite families of discrete conditional probability distributions, represented by two stochastic matrices P and Q, is obtained. Moreover, the possible extensions of the assessment (P,Q) to the marginal distributions are examined and explicit formulas for them are given in some special cases. Finally, a general algorithm to check coherence of (P,Q) and to derive its extensions is proposed.

#### 1. Introduction.

Given two sets of integers  $I = \{1, 2, ..., l\}$  and  $J = \{1, 2, ..., m\}$ , let  $\{A_i, i \in I\}$  and  $\{B_j, j \in J\}$  be two finite partitions of the certain event  $\Omega$  and consider the two families of conditional events  $\mathcal{F}_1 = \{B_j/A_i, i \in I, j \in J\}$  and  $\mathcal{F}_2 = \{A_i/B_j, i \in I, j \in J\}$ . Given an  $l \times m$  stochastic matrix  $P = \{p_{ij}\}$ , we can introduce a coherent probability assessment on  $\mathcal{F}_1$  assigning, for each i, j, the probability  $p_{ij}$  to the event  $B_j/A_i$ . Analogously, using an  $m \times l$  stochastic matrix  $Q = \{q_{ji}\}$ , we can introduce a coherent probability assessment on  $\mathcal{F}_2$ .

Entrato in Redazione il 4 maggio 1995.

Key words: Conditional events, Coherence principle, Discrete probability distributions, Stochastic matrices.

In this way, a probability assessment, denoted by (P, Q), is defined on  $\mathcal{F}_1 \cup \mathcal{F}_2$ . However, we observe that (P, Q) is not generally coherent.

A concrete interpretation of the above partitions is obtained considering two random variables X and Y, whose values are respectively  $\{x_1, \ldots, x_l\}$  and  $\{y_1, \ldots, y_m\}$ , and defining  $A_i = (X = x_i)$  and  $B_j = (Y = y_j)(i \in I, j \in J)$ . The problem of finding the marginal distributions of X and Y, i.e. the probabilities of the unconditional events  $A_i$ ,  $B_j$ , from the conditional ones represented by P and Q, has been considered in [20].

The assessment (P, Q) has been considered as representing the knowledge base of an expert system in [11], where, using the coherence principle of de Finetti [7], the probabilistic consistency and the extension to a complete probabilistic model have been analyzed. Some results on the characterization of coherent conditional probabilities are given in [4] and [5]. The probabilistic treatment of uncertainty in artificial intelligence, based on de Finetti's approach, has been adopted in many papers (see for example [3], [10], [19]).

In this paper we reconsider the problem introduced in [11] obtaining further results on coherence of (P,Q) and on its extensions to marginal distributions. In Section 2 some preliminary results on coherence are given. We first recall a general result, given in [18], on coherence of a conditional probability, based on a condition introduced in [6], and we give a different proof of it. Then, applying this result, in Section 3 we obtain a necessary and sufficient condition for coherence of (P,Q). In Section 4 the special case of P and Q each one having at least one positive column is considered and an explicit condition for the existence of a coherent extension to the marginal distributions is given. The general case is analyzed in Section 5 using a suitable partition of the matrices P and Q. In Section 5.1 we briefly review some results on coherence of (P,Q) obtained in [11]. Finally, a general algorithm to check coherence of (P,Q) and to derive the marginal distributions is proposed in Sections 5.2 and 5.3.

#### 2. Some results on coherence.

In this Section some results on coherence of conditional probability assessments are considered.

Denote by  $\mathcal{K}$  and arbitrary family of conditional events and by  $\mathcal{P}$  a real function defined on  $\mathcal{K}$ . We denote by  $\Omega$  and  $\emptyset$  the certain and the impossible events and by  $\overline{E}$  the contrary of the event E. Moreover, the symbol AB denotes the logical product of the events A and B.

Given *n* conditional events  $E_1/H_1, \ldots, E_n/H_n$  belonging to  $\mathcal{K}$ , denote by  $C_1, \ldots, C_s$  the atoms generated by  $E_i, H_i$   $(i = 1, 2, \ldots, n)$  and contained in

 $H_1 \cup ... \cup H_n$ . Letting  $p_i = \mathcal{P}(E_i/H_i)$  (i = 1, ..., n) we associate the following loss function to the point  $P = (p_1, p_2, ..., p_n)$ 

$$\mathcal{L} = \sum_{1}^{n} {}_{i}H_{i}(E_{i} - p_{i})^{2},$$

where the same symbol denotes the event and its indicator. Moreover, let  $L_1, \ldots, L_s$  be the values of  $\mathcal{L}$  corresponding to the atoms  $C_1, \ldots, C_s$ . Based on de Finetti's penalty criterion (see [7]), in [8] the real function  $\mathcal{P}$  is defined coherent if, for every  $n = 1, 2, \ldots$  and  $\mathcal{F} = \{E_1/H_1, \ldots, E_n/H_n\} \subseteq \mathcal{K}$ , there does not exist a point  $(p_1^*, p_2^*, \ldots, p_n^*) \neq (p_1, p_2, \ldots, p_n)$  such that  $L_h^* \leq L_h$  for every subscript h, with  $L_k^* < L_k$  for at least a subscript k. Moreover, in the same paper the generalized atoms  $Q_1, \ldots, Q_s$  relative to  $\mathcal{F}$  and P are

(1) 
$$\alpha_{hi} = \begin{cases} 1, & \text{if } C_h \subseteq E_i H_i \\ 0, & \text{if } C_h \subseteq \overline{E}_i H_i \end{cases}, \quad h = 1, 2, \dots, s,$$
$$p_i, & \text{if } C_h \subseteq \overline{H}_i \end{cases}$$

introduced, defining  $Q_h = (\alpha_{h1}, \dots, \alpha_{hn})$ , where

and the following result is proved:

**Proposition 1.** The real function  $\mathcal{P}$  defined on the class of conditional events  $\mathcal{K}$  is coherent if and only if, for every  $n = 1, 2, \ldots$  and

$$\mathcal{F} = \{E_1/H_1, \ldots, E_n/H_n\} \subseteq \mathcal{K},$$

the point  $P = (p_1, p_2, ..., p_n)$  is a mixture of the generalized atoms relative to  $\mathcal{F}$  and P.

Coherence can be alternatively defined by means of the betting criterion (see [7], [13], [14], [15], [16], [17], [21]) and it can be shown that the two definitions are equivalent (see [8]).

We observe that the family  $\mathcal{K}$  considered in Proposition 1 is not required to have any algebraic property. Alternative ways to check coherence when the family  $\mathcal{K}$  has some particular structure can be given.

We recall that a real function  $\mathcal{P}$  defined on  $\mathcal{E} \times \mathcal{H}$ , where  $\mathcal{E}$  is an algebra of events and  $\mathcal{H} \subset \mathcal{E}$  is a non empty family of events not containing  $\emptyset$ , is named a conditional probability if the following properties are satisfied

- i)  $\mathcal{P}(E/H) \ge 0$  and  $\mathcal{P}(H/H) = 1, E \in \mathcal{E}, H \in \mathcal{H}$ ;
- ii)  $\mathcal{P}(A \cup B/H) = \mathcal{P}(A/H) + \mathcal{P}(B/H)$ , if  $AB = \emptyset$ ,  $A, B \in \mathcal{E}, H \in \mathcal{H}$ ,
- iii)  $\mathcal{P}(AB/H) = \mathcal{P}(A/BH)\mathcal{P}(B/H), A, B \in \mathcal{E}, H, BH \in \mathcal{H}.$

As well known, conditions i), ii) and iii) are not sufficient for coherence of  $\mathcal{P}$ .

**Remark.** Using the betting criterion Holzer ([14]) has proved that, when  $\mathcal{H}$  is an additive class, the conditional probability  $\mathcal{P}$  is coherent.

Given a conditional probability  $\mathcal{P}$ , the following condition, which is related to coherence of  $\mathcal{P}$ , has been introduced in [6] by Császár:

(2) 
$$\prod_{i=1}^{n} \mathcal{P}(E_i/H_i) = \prod_{i=1}^{n} \mathcal{P}(E_i/H_{i+1}),$$

where  $E_i \in \mathcal{E}$ ,  $H_i \in \mathcal{H}$ ,  $E_i \subseteq H_i H_{i+1}$ , and  $H_{n+1} = H_1$ .

In the same paper (Theorem 8.5) Császár has also proved the following proposition:

**Proposition 2.**  $\mathcal{P}$  satisfies (2) if and only if there exists a conditional probability  $\mathcal{P}^*$ , extension of  $\mathcal{P}$ , defined on  $\mathcal{E} \times \mathcal{H}^*$ , where  $\mathcal{H}^*$  is an additive class containing  $\mathcal{H}$ .

**Remark.** Proposition 2 has been proved assuming that  $\mathcal{E}$  is a  $\sigma$ -algebra and  $\mathcal{P}$  is  $\sigma$ -additive, but, as explicitly noted in Császár's paper (Sec. 9, p. 360), the same result holds in a finitely additive framework.

The relationship between coherence of  $\mathcal{P}$  and Császár's condition has been considered in [18] and can be expressed by the following theorem:

**Theorem 3.** A conditional probability  $\mathcal{P}$  defined on  $\mathcal{E} \times \mathcal{H}$ , where  $\mathcal{E}$  is an algebra of events and  $\mathcal{H} \subset \mathcal{E}$  is a non empty family of events not containing  $\emptyset$ , is coherent if and only if, for each n, condition (2) is satisfied.

We will give a proof of Theorem 3 based on the concept of generalized atom (for a different proof see [1]).

*Proof.* Assume that  $\mathcal{P}$  is coherent. Given the events  $E_i/H_i$ ,  $E_i/H_{i+1}$ , with  $E_i \in \mathcal{E}$ ,  $H_i \in \mathcal{H}$ ,  $E_i \subseteq H_iH_{i+1}$  (i = 1, 2, ..., n) and  $H_{n+1} = H_1$ , let  $\mathcal{P}(E_i/H_i) = p_i$ ,  $\mathcal{P}(E_i/H_{i+1}) = q_i$  (i = 1, 2, ..., n) and denote by  $Q_1, ..., Q_s$  the generalized atoms relative to  $E_i/H_i$ ,  $E_i/H_{i+1}$  and to  $p_i$ ,  $q_i$  (i = 1, 2, ..., n).

From (1) we have  $Q_h = (a_{h1}, ..., a_{hn}, b_{h1}, ..., b_{hn})$ , where, for i = 1, 2, ..., n, and h = 1, 2, ..., s, we put

$$a_{hi} = \begin{cases} 1, & \text{if} \quad C_h \subseteq E_i H_i \\ 0, & \text{if} \quad C_h \subseteq \overline{E}_i H_i \\ p_i, & \text{if} \quad C_h \subseteq \overline{H}_i \end{cases},$$

$$b_{hi} = \begin{cases} 1, & \text{if} \quad C_h \subseteq E_i H_{i+1} \\ 0, & \text{if} \quad C_h \subseteq \overline{E}_i H_{i+1} \\ q_i & \text{if} \quad C_h \subseteq \overline{H}_{i+1} \end{cases}.$$

From Proposition 1, coherence of  $\mathcal{P}$  implies that the point

$$(p_1,\ldots,p_n,q_1,\ldots,q_n)$$

is a mixture of the points  $Q_1, \ldots, Q_s$ , that is:

$$p_i = \sum_{h=1}^{s} \lambda_h a_{hi}, \quad q_i = \sum_{h=1}^{s} \lambda_h b_{hi}, \quad (i = 1, 2, ..., n)$$

where  $\lambda_h \geq 0$  and  $\sum_{h=1}^{s} \lambda_h = 1$ .

Given an event A we denote by I(A) the set of integers h such that  $C_h \subseteq A$ . Since  $E_i \subseteq H_i H_{i+1}$ , we have:

$$\sum_{h\in I(E_i)}\lambda_h=\sum_{h\in I(E_iH_i)}\lambda_h=p_i-p_i\sum_{h\in I(\overline{H}_i)}\lambda_h=p_i\sum_{h\in I(H_i)}\lambda_h,$$

and

$$\sum_{h\in I(E_i)} \lambda_h = \sum_{h\in I(E_i H_{i+1})} \lambda_h = q_i - q_i \sum_{h\in I(\overline{H}_{i+1})} \lambda_h = q_i \sum_{h\in I(H_{i+1})} \lambda_h,$$

therefore  $p_i S_i = q_i S_{i+1}$ , where  $S_i = \sum_{h \in I(H_i)} \lambda_h$  (i = 1, 2, ..., n) and

$$\prod_i p_i S_i = \prod_i q_i S_{i+1}.$$

We observe that  $\prod_i S_i = \prod_i S_{i+1}$  and, since  $\sum_i S_i \ge \sum_{h=1}^s \lambda_h = 1$ , there exists a subscript i' such that  $S_{i'} > 0$ . If  $S_i > 0$  for each i, then  $\prod_i p_i = \prod_i q_i$ , that is (2).

If instead  $S_i = 0$  for some i, then there exist two integers j and k, with  $j, k \in \{1, 2, ..., n\}$ , such that  $S_j = 0$ ,  $S_{j+1} > 0$  and  $S_k > 0$ ,  $S_{k+1} = 0$ . Therefore, since  $p_j S_j = q_j S_{j+1} = 0$  and  $p_k S_k = q_k S_{k+1} = 0$ , we have  $q_j = p_k = 0$ , which implies (2).

Conversely, assume that (2) is satisfied. Then  $\mathcal{P}$  can be extended to  $\mathcal{P}^*$  defined on  $\mathcal{E} \times \mathcal{H}^*$ , where  $\mathcal{H}^*$  is an additive class (see Proposition 2). We now prove that, for each n, given the family  $\mathcal{F} = \{E_1/H_1, E_2/H_2, \ldots, E_n/H_n\}, E_i \in \mathcal{E}, H_i \in \mathcal{H}^*$ , and letting  $\mathcal{P}^*(E_i/H_i) = p_i$ , the point  $P = (p_1, p_2, \ldots, p_n)$  is a mixture of the generalized atoms relative to  $\mathcal{F}$  and P.

Consider the atoms  $C_1, \ldots, C_s$  generated by  $E_i, H_i$   $(i = 1, 2, \ldots, n)$  and

contained in  $H_0 = H_1 \cup \ldots \cup H_n$  and the generalized atoms  $Q_h = (a_{h1}, \ldots, a_{hn})$   $(h = 1, 2, \ldots, s)$  relative to  $\mathcal{F}$  and P. Then, letting  $\lambda_h = \mathcal{P}^*(C_h/H_0)$ ,  $(h = 1, 2, \ldots, s)$  it is, for each i

$$\mathcal{P}^*(E_i H_i/H_0) = \sum_{h \in I(E_i H_i)} \lambda_h, \quad \mathcal{P}^*(H_i/H_0) = \sum_{h \in I(H_i)} \lambda_h.$$

From the relation

$$\mathcal{P}^*(E_i H_i / H_0) = \mathcal{P}^*(E_i / H_i) \mathcal{P}^*(H_i / H_0) = p_i \mathcal{P}^*(H_i / H_0),$$

using (1) it follows that  $p_i = \sum_{h=1}^{s} \lambda_h a_{hi}$ , i = 1, 2, ..., n. Therefore, from Proposition 1,  $\mathcal{P}^*$  is coherent. Hence its restriction  $\mathcal{P}$  is coherent too.

From Theorem 3, a useful result on coherence of  $\mathcal{P}$  which will be used in the next Sections is the following (see [2], Corollary 1.3):

Corollary 4. Let  $\mathcal{H} = \Pi_1 \cup \{\Omega\} \cup \Pi_2$ , where  $\Pi_1$  and  $\Pi_2$  are two partitions of  $\Omega$  contained in  $\mathcal{E}$ . Then, a conditional probability  $\mathcal{P}$ , defined on  $\mathcal{E} \times \mathcal{H}$ , is coherent if and only if condition (2) is satisfied for the  $H_i$  's such that  $\mathcal{P}(H_i) = 0$ , with  $H_i \in \Pi_r$  and  $H_{i+1} \in \Pi_s$   $(r, s = 1, 2, r \neq s)$ .

# 3. Coherence and extensions of (P, Q).

Before considering the problem of the extension of (P, Q), from Theorem 3 a necessary and sufficient condition for its coherence is given.

**Corollary 5.** Given the  $(l \times m \text{ and } m \times l)$  stochastic matrices P and Q, the conditional probability assessment (P, Q), defined on  $\mathcal{F}_1 \cup \mathcal{F}_2$ , is coherent if and only if

$$(3) P_{i_1j_1}q_{j_1i_2}\cdots p_{i_hj_h}q_{j_hi_{h+1}}\cdots p_{i_tj_t}q_{j_ti_1}=p_{i_1j_t}q_{j_ti_t}\cdots p_{i_{h+1}j_h}q_{j_hi_h}\cdots p_{i_2j_1}q_{j_1i_1},$$

where  $t \leq \min(l, m)$ ,  $i_k \in I$ ,  $j_k \in J$  (k = 1, 2, ..., t) and  $i_r \neq i_h$ ,  $j_r \neq j_h$  if  $r \neq h$ .

*Proof.* Corollary 5 follows from Theorem 3 extending the assessment (P, Q) to a conditional probability P on  $E \times H$ , where E is the algebra generated by the family  $H = \{A_i, B_j, i \in I, j \in J\}$ , and showing that condition (2) reduces to (3).

The extension of (P, Q) to a conditional probability  $\mathcal{P}$  is obtained observing that the atoms of  $\mathcal{E}$  are the events  $A_i B_j$   $(i \in I, j \in J)$ . Then, the probability

 $\mathcal{P}(E/H)$ , for each  $E/H \in \mathcal{E} \times \mathcal{H}$ , is defined using the quantities  $\mathcal{P}(A_i B_j/H)$  and the additive property, where

$$\mathcal{P}(A_i B_j / H) = \begin{cases} 0, & \text{if} \quad H \neq A_i, B_j \\ p_{ij}, & \text{if} \quad H = A_i \\ q_{ji}, & \text{if} \quad H = B_j \end{cases}.$$

It can be verified that  $\mathcal{P}$  satisfies the properties i), ii) and iii) of a conditional probability (see also [9], Example 8).

Consider  $E_i \in \mathcal{E}$ ,  $H_i \in \mathcal{H}$ ,  $E_i \subseteq H_iH_{i+1}$  ( $i=1,\ldots,n$ ) with  $H_{n+1}=H_1$ . To avoid the trivial cases of equal factors in both sides of (2) we consider  $E_i$ ,  $H_i$ ,  $H_{i+1}$  ( $i=1,\ldots,n$ ) such that  $E_i \neq \emptyset$  (which implies that  $H_iH_{i+1} \neq \emptyset$ ) and  $H_i \neq H_{i+1}$ . Then, the possible choices for  $E_i$ ,  $H_i$ ,  $H_{i+1}$  are  $H_i = A_h$  ( $= B_k$ ),  $H_{i+1} = B_k$  ( $= A_h$ ) for some (h, k), with  $h \in I$ ,  $k \in J$ , and  $E_i = H_iH_{i+1} = A_hB_k$  ( $i=1,\ldots,n$ ). Moreover, we observe that, if the integer n is odd, being  $H_{n+1} = H_1$ , we cannot have  $H_iH_{i+1} = A_hB_k$  for i=n. Therefore, the relevant cases for condition (2) are obtained when n is even and  $E_i$ ,  $H_i$ ,  $H_{i+1}$  are such that  $\mathcal{P}(E_i/H_i) = p_{hk} (=q_{kh})$ ,  $\mathcal{P}(E_i/H_{i+1}) = q_{kh} (=p_{hk})$ , for some (h, k), with  $h \in I$ ,  $k \in J$ , ( $i=1,\ldots,n$ ). So that in the non trivial cases condition (2), with n=2t, becomes (3). Moreover, if (3) is satisfied for every t, with  $i_k \neq i_h$ ,  $j_k \neq j_h$  for  $k \neq h$ , then (3) is satisfied also when  $i_k = i_h$  or  $j_k = j_h$  for some (h, k). In fact, if for instance  $i_k \neq i_h$ ,  $j_k \neq j_h$  for  $k \neq h$ , with the only exception  $i_h = i_1$  for a subscript  $\bar{h} \in \{2,\ldots,t\}$ , then (3) follows from the equalities

$$p_{i_1j_1}q_{j_1i_2}\cdots p_{i_{\bar{h}-1}}j_{\bar{h}-1}q_{j_{\bar{h}-1}}i_1 = p_{i_1j_{\bar{h}-1}}q_{j_{\bar{h}-1}}i_{\bar{h}-1}\cdots p_{i_2j_1}q_{j_1i_1}$$
$$p_{i_1j_{\bar{h}}}q_{j_{\bar{h}}i_{\bar{h}+1}}\cdots p_{i_rj_r}q_{j_ri_1} = p_{i_1j_r}q_{j_ri_r}\cdots p_{i_{\bar{h}+1}j_{\bar{h}}}q_{j_{\bar{h}}i_1}.$$

Finally, we note that it is sufficient to consider (3) for  $t \leq \min(l, m)$ . In fact, observing that  $t > \min(l, m)$  implies  $i_k = i_h$  or  $j_k = j_h$  for at least one pair of subscripts (h, k), it can be shown, by a similar argument as the previous one, that the equality (3), when  $t > \min(l, m)$ , can be obtained from suitable equalities (3), with  $t \leq \min(l, m)$ . Therefore condition (2) is equivalent to (3) and Corollary 5 is proved.

**Remark.** If the assessment (P, Q) is coherent, then, as shown in the proof of Corollary 5, (P, Q) can be extended to a conditional probability P on  $E \times H$  satisfying condition (2). Now, from Proposition 2, P can be extended on  $E \times H^*$ , where  $H^*$  is an additive class containing the certain event

$$\Omega = A_1 \cup A_2 \cup \ldots \cup A_l = B_1 \cup B_2 \cup \ldots \cup B_m.$$

Therefore,  $\mathcal{P}$  can be extended to the events  $A_i$ ,  $B_j$   $(i \in I, j \in J)$  providing marginal distributions of X and Y. Denoting the extension by the same symbol  $\mathcal{P}$ , we define  $f_i = \mathcal{P}(A_i)$  and  $g_j = \mathcal{P}(B_j)$   $(i \in I, j \in J)$ . We observe that the marginal distributions  $\{f_i\}$ ,  $\{g_j\}$  satisfy the following linear system:

(4) 
$$p_{ij}f_i = q_{ji}g_j, \ f_i \ge 0, \ g_j \ge 0, \ \sum f_i = \sum g_j = 1 \ (i \in I, j \in J),$$

so that compatibility of (4) is a necessary condition for coherence of (P, Q). Conversely, if (4) admits a solution  $\{f_i\}$ ,  $\{g_j\}$ , then we can extend (P, Q) to a conditional probability  $\mathcal{P}$  on  $\mathcal{E} \times \mathcal{H}$ , where  $\mathcal{E}$  is the algebra generated by the family  $\mathcal{H} = \{\Omega, A_i, B_j, i \in I, j \in J\}$ , defining  $\mathcal{P}(A_i) = f_i$  and  $\mathcal{P}(B_j) = g_j$ . Then, in order to verify the coherence of (P, Q), we can apply Corollary 4, with  $\Pi_1 = \{A_i, i \in I\}$  and  $\Pi_2 = \{B_j, j \in J\}$ , which amounts to verify (3) for  $i_h, j_k$  such that  $f_{i_h} = g_{j_k} = 0$ .

From the previous remark the following result is immediately obtained.

**Proposition 6.** If there exists a strictly positive solution  $\{f_i\}$ ,  $\{g_j\}$  of system (4), then (P, Q) is coherent.

# 4. A special case for P and Q.

Assume that P and Q satisfy the following property:

**Property 7.** There exist two indices  $h_0$ ,  $k_0$  such that  $p_{ih_0} > 0$ ,  $q_{jk_0} > 0$   $(i \in I, j \in J)$ . Then we have the following theorem:

**Theorem 8.** If P and Q verify Property 7, then the system (4) is compatible if and only if the following condition is satisfied:

$$(5) p_{ij}q_{jk_0}p_{k_0h_0}q_{h_0i} = p_{ih_0}q_{h_0k_0}p_{k_0j}q_{ji} (i \in I, j \in J).$$

Moreover, if system (4) is compatible, then it has a unique solution given by

(6) 
$$f_i^* = (q_{h_0i}/p_{ih_0}) \left(\sum_r q_{h_0r}/p_{rh_0}\right)^{-1},$$
$$g_j^* = (p_{k_0j}/q_{jk_0}) \left(\sum_r p_{k_0r}/q_{rk_0}\right)^{-1} \quad (i \in I, j \in J).$$

*Proof.* Let  $\{f_i, g_j, i \in I, j \in I\}$  be a solution of (4). Then, since

$$g_{h_0} = \sum_i p_{ih_0} f_i, f_{k_0} = \sum_j q_{jk_0} g_j,$$

from Property 7 we have

$$g_{h_0} > 0$$
,  $f_{k_0} > 0$ .

Therefore

$$p_{ij}(q_{h_0i}/p_{ih_0}) = p_{ij}(f_i/g_{h_0}) = q_{ji}(q_{h_0k_0}/p_{k_0h_0})(p_{k_0j}/q_{jk_0}),$$

that is (5).

Conversely, the sufficient part follows by verifying that (6) is a solution of system (4). In fact, from (5) we have

(7) 
$$q_{h_0u}/p_{uh_0} = \sum_{v} q_{vu}(p_{k_0v} q_{h_0k_0}/p_{k_0h_0} q_{vk_0}) \quad (u \in I)$$

and, using (5) and (6), from (7), for each i, j, it follows

$$p_{ij} f_i^* = q_{ji} (p_{k_0 j} q_{h_0 k_0} / p_{k_0 h_0} q_{j k_0}) \left( \sum_{u} q_{h_0 u} / p_{u h_0} \right)^{-1} =$$

$$= q_{ji} (p_{k_0 j} / q_{j k_0}) \left( \sum_{u} \sum_{v} q_{v u} p_{k_0 v} / q_{v k_0} \right)^{-1} = q_{ji} g_j^*,$$

that is the values  $f_i^*$ ,  $g_j^*$   $(i \in I, j \in J)$  are a solution of (4). Moreover, from (6), we have

(8) 
$$f_i^* > 0 \Longleftrightarrow q_{h_0i} > 0$$
 and  $g_j^* > 0 \Longleftrightarrow p_{k_0j} > 0$   $(i \in I, j \in J)$ .

To show the uniqueness of solution (6), assume that  $\{f_i, g_j, i \in I, j \in J\}$  is a solution of (4), so that

$$(9) p_{ih_0}f_i = q_{h_0i}g_{h_0} (i \in I).$$

From (8) and (9) we have

$$g_{h_0}/f_{k_0} = p_{k_0h_0}/q_{h_0k_0} = g_{h_0}^*/f_{k_0}^*$$

and

$$p_{ih_0}f_i = q_{h_0i}f_{k_0}(g_{h_0}^*/f_{k_0}^*) = p_{ih_0}f_i^*(f_{k_0}/f_{k_0}^*),$$

from which it follows that

$$f_i = f_i^* (f_{k_0}/f_{k_0}^*) \quad (i \in I).$$

Then, since  $\sum_i f_i = \sum_i f_i^* = 1$ , we obtain  $f_i = f_i^*$   $(i \in I)$ . The same argument can be applied to show that  $g_j = g_j^*$   $(j \in J)$ .

Remark. Property 7, adapted to the continuous case, has been used in [12] to uniquely determine the joint density of a random vector (X, Y) from the conditional densities.

**Remark.** If condition (5) is satisfied, system (4) is compatible. However, as noted in the remark at the end of Section 3, compatibility of system (4) does not imply coherence of (P, Q). For an example, see [11]. By the same remark and from (8) we have that in order to check coherence it is sufficient to verify (3) for  $i_h$ ,  $j_k$  such that  $q_{h_0i_h} = p_{k_0j_k} = 0$ .

Remark. Note that in the important case in which P and Q are strictly positive, if condition (5) is satisfied, the solution (6) of system (4) is strictly positive and (P, Q) is coherent. So that, in this case, compatibility of system (4) is equivalent to coherence of (P, Q).

### 5. The general case.

Given two arbitrary stochastic matrices P and Q, consider the subsets  $I \setminus I_0$  and  $J \setminus J_0$  defined as

$$I \setminus I_0 = \{i \in I : \exists j \text{ such that } q_{ji} = 0, p_{ij} > 0\}$$
 and 
$$J \setminus J_0 = \{j \in J : \exists i \text{ such that } p_{ij} = 0, q_{ji} > 0\}.$$

Then

(10) 
$$p_{ij} = 0$$
 if and only if  $q_{ji} = 0$   $(i \in I_0, j \in J_0)$ .

The system (4) can be represented in the following way:

$$(4a) p_{ij} f_i = q_{ji} g_j, \quad i \in I \setminus I_0, \quad j \in J \setminus J_0,$$

$$(4b) p_{ij} f_i = q_{ji} g_j, \quad i \in I \setminus I_0, \ j \in J_0,$$

$$(4c) p_{ij} f_i = q_{ji} g_j, \quad i \in I_0, \ j \in J \setminus J_0,$$

(4d) 
$$p_{ij} f_i = q_{ji} g_j, \quad i \in I_0, \ j \in J_0,$$
  $f_i \ge 0, \ g_i \ge 0, \ \sum f_i = \sum g_i = 1,$ 

and from definition of sets  $I \setminus I_0$  and  $J \setminus J_0$  the following necessary condition for the compatibility of (4a), (4b) and (4c) is obtained

(11) 
$$f_i = 0, \quad i \in I \setminus I_0; \quad g_j = 0, \quad j \in J \setminus J_0.$$

Moreover, given a solution  $\{f_i, g_j, i \in I, j \in J\}$  of system (4), from (11) it follows that the set of values  $\{f_i, g_j, i \in I_0, j \in J_0\}$  is a solution of the following system

(12) 
$$p_{ij} f_i = q_{ji} g_j, f_i \ge 0, g_j \ge 0, \sum f_i = \sum g_j = 1 \quad (i \in I_0, j \in J_0).$$

Based on the above representation, the following submatrices can be introduced

$$P_a = \{p_{ij}\}, \quad Q_a = \{q_{ji}\}, \quad i \in I \setminus I_0, \ j \in J \setminus J_0,$$
 $P_b = \{p_{ij}\}, \quad Q_b = \{q_{ji}\}, \quad i \in I \setminus I_0, \ j \in J_0,$ 
 $P_c = \{p_{ij}\}, \quad Q_c = \{q_{ji}\}, \quad i \in I_0, \ j \in J \setminus J_0,$ 
 $P_d = \{p_{ij}\}, \quad Q_d = \{q_{ji}\}, \quad i \in I_0, \ j \in J_0.$ 

Without loss of generality, we can assume that

$$P = \begin{bmatrix} P_a & P_b \\ P_c & P_d \end{bmatrix}, \quad Q = \begin{bmatrix} Q_a & Q_c \\ Q_b & Q_d \end{bmatrix}.$$

In a previous paper (see [11], Section 4.3.1) the relationship between coherence of (P, Q) and compatibility of system (4) has been analyzed in the special case  $I_0 = I$ ,  $J_0 = J$ . In Section 5.1 we briefly review this special case. In Section 5.2, assuming  $P_d$  and  $Q_d$  stochastic, the general case  $I_0 \subset I$  and/or  $J_0 \subset J$  is analyzed using the results of Section 5.1. Finally, the case of  $P_d$  and  $Q_d$  not stochastic is considered in Section 5.3.

## **5.1.** *P* and *Q* such that $I_0 = I$ , $J_0 = J$ .

Before analyzing the case  $I_0 = I$ ,  $J_0 = J$  we recall the definition of connected matrix (see [20]): an  $l \times m$  matrix  $W = \{w_{ij}\}$  is connected if for every pair of distinct integers  $j_1$ ,  $j_{q+1}$  in the set J there exists a subset  $S \subset I \times J$ , where

$$S = \{(i_h, j_h), h = 1, \dots, q, i_r \neq i_s \text{ if } r \neq s, \text{ with } r, s = 1, \dots, q,$$
  
and  $j_r \neq j_s \text{ if } r \neq s, \text{ with } r, s = 1, \dots, q + 1\},$ 

such that

$$w_{i_1j_1}w_{i_1j_2}w_{i_2j_2}w_{i_2j_3}\cdots w_{i_qj_q}w_{i_qj_{q+1}}>0.$$

Therefore a matrix W is connected if there are no row and column permutations which change W to the form

$$\begin{bmatrix} A_{rs} & \theta_{r,m-s} \\ \theta_{l-r,s} & B_{l-r,m-s} \end{bmatrix},$$

for some r, s, such that  $1 \le r < l$  and  $1 \le s < m$ . Two cases can be considered:

(i) P (and therefore Q) connected

We observe that in this case (4) is equivalent to the following system  $\Sigma$ 

$$p_{ij} f_i = q_{ji} g_j, \ f_i > 0, \ g_j > 0, \ \sum f_i = \sum g_j = 1 \ (i \in I, \ j \in J).$$

In fact, any solution of  $\Sigma$  is a solution of (4) too. Conversely, let  $\{f_i, g_j, i \in I, j \in J\}$  be a solution of system (4). Now, given two distinct integers  $j_1, j_{q+1}$  in the set J, from the definition of connected matrix, there exists a subset  $S \subset I \times J$ , where

$$S = \{(i_h, j_h), h = 1, \dots, q, i_r \neq i_s \text{ if } r \neq s, \text{ with } r, s = 1, \dots, q,$$
  
and  $j_r \neq j_s \text{ if } r \neq s, \text{ with } r, s = 1, \dots, q + 1\},$ 

such that

$$p_{i_1j_1}p_{i_1j_2}p_{i_2j_2}p_{i_2j_3}\cdots p_{i_aj_a}p_{i_aj_{a+1}}>0.$$

Then, from (10), we have

$$q_{j_1i_1}q_{j_2i_1}q_{j_2i_2}q_{j_3i_2}\cdots q_{j_qi_q}q_{j_{q+1}i_q}>0,$$

and hence

$$g_{j_1} = (p_{i_1j_1}/q_{j_1i_1})f_{i_1}, f_{i_1} = (q_{j_2i_1}/p_{i_1j_2})g_{j_2}, g_{j_2} = (p_{i_2j_2}/q_{j_2i_2})f_{i_2},$$

$$f_{i_2} = (q_{j_3i_2}/p_{i_2j_3})g_{j_3}, \cdots, g_{j_q} = (p_{i_qj_q}/q_{j_qi_q})f_{i_q}, f_{i_q} = (q_{j_{q+1}i_q}/p_{i_qj_{q+1}})g_{j_{q+1}}.$$

Then  $g_{j_1} = \lambda g_{j_{q+1}}$ , where

$$\lambda = (p_{i_1j_1}q_{j_2i_1}p_{i_2j_2}q_{j_3i_2}\cdots p_{i_qj_q}q_{j_{q+1}i_q})/(q_{j_1i_1}p_{i_1j_2}q_{j_2i_2}p_{i_2j_3}\cdots q_{j_qi_q}p_{i_qj_{q+1}}).$$

Since  $j_1, j_{q+1}$  are arbitrarily chosen and  $\sum g_j = 1$ , it follows that  $g_j > 0, j \in J$ . Applying the same procedure to Q we also obtain  $f_i > 0, i \in I$ . Therefore  $\{f_i, g_j, i \in I, j \in J\}$  is a solution of system  $\Sigma$ .

Let N(P) be the number of positive elements of P. Since P is connected, then (see [20])

- i)  $N(P) \ge l + m 1$ ;
- ii) it is possible to select at least one set  $D \subset I \times J$  such that the matrix U with elements  $u_{ij} = p_{ij}$ , or 0, according to whether  $(i, j) \in D$ , or  $(i, j) \in I \times J \setminus D$ , is connected and N(U) = l + m 1.

In order to find conditions on compatibility of  $\Sigma$ , let V be a matrix with elements  $v_{ji} = q_{ji}$ , or 0 according to whether  $(i, j) \in D$ , or  $(i, j) \in I \times J \setminus D$ . Then define the matrices  $T = (t_{ij}) = (UV)^{\gamma-1}U$ ,  $Z = (z_{ji}) = (VU)^{\gamma-1}V$ , where  $\gamma$  is the smallest integer such that T and Z are strictly positive. In [20] it has been shown that:

- (a)  $\gamma \leq \min(l, m)$ ;
- (b)  $\Sigma$  is compatible if and only if the following condition, which does not depend on the choice of the set D, is satisfied:

(13) 
$$t_{ij}/z_{ji} = p_{ij}/q_{ji}, \quad \forall (i,j): p_{ij} > 0;$$

(c) if  $\Sigma$  is compatible, the (unique and positive) solution is

$$f_i^* = \left(\sum_h t_{ih}/z_{hi}\right)^{-1}, \quad g_j^* = \left(\sum_k z_{jk}/t_{kj}\right)^{-1}.$$

We have shown that, if system (4) is compatible, its solution is strictly positive. Therefore, by Proposition 6, the compatibility of system (4) implies the coherence of (P, Q). Summarizing, in order to check the coherence of the assessment (P, Q) we have to define two connected matrices U and V, to compute T and Z and to verify condition (13).

## (ii) P (and therefore Q) not connected

In this case there exists an integer k, with  $2 \le k \le \min(l, m)$ , such that, under a suitable row and column permutation, the following partition of P and Q can be obtained

$$egin{bmatrix} P_1 & 0 & \dots & 0 \ 0 & P_2 & \dots & 0 \ \dots & \dots & \dots & \dots \ 0 & 0 & \dots & P_k \end{bmatrix} \quad , \quad egin{bmatrix} Q_1 & 0 & \dots & 0 \ 0 & Q_2 & \dots & 0 \ \dots & \dots & \dots \ 0 & 0 & \dots & Q_k \end{bmatrix} ,$$

where  $P_1, \ldots, P_k, Q_1, \ldots, Q_k$  are connected stochastic matrices. We denote by  $I_h \times J_h$  the set of pairs of integers (i, j) such that  $p_{ij}$   $(q_{ji})$  is an element of  $P_h$   $(Q_h)$   $(h = 1, \ldots, k)$ . We have the following result

**Proposition 9.** The assessment (P, Q) is coherent if and only if each assessment  $(P_h, Q_h)$  (h = 1, ..., k) is coherent.

The proof of Proposition 9 has been given, for k = 2, in a previous paper (see [11], Proposition 7). The general case is proved by iteratively applying Proposition 7 of the quoted paper.

For each h, to check coherence of  $(P_h, Q_h)$  we can apply the procedure described in case (i) to analyze the compatibility of the following system:

$$p_{ij} f_i = q_{ji} g_j, \ \sum f_i = \sum g_j = 1, \ f_i \ge 0, \ g_j \ge 0 \quad (i \in I_h, j \in J_h).$$

Assume that, for each h, the above system is compatible and denote by  $(\phi_h, \gamma_h)$  its solution, where  $\phi_h, \gamma_h$  are strictly positive vectors. Then (P, Q) is coherent and, for each  $(\theta_1, \ldots, \theta_k)$ , with  $\theta_h \in [0, 1]$   $(h = 1, \ldots, k)$  and  $\sum \theta_h = 1$ , a solution  $(\phi, \gamma)$  of (4), where

$$\phi = (\theta_1 \phi_1, \dots, \theta_k \phi_k), \quad \gamma = (\theta_1 \gamma_1, \dots, \theta_k \gamma_k),$$

is obtained. Each pair of vectors  $(\phi, \gamma)$  provides a coherent extension of the assessment (P, Q) to the marginal distributions.

We observe that if (at least) one of the above systems is not compatible, then (P, Q) is not coherent.

Moreover, if (at least) one of the above systems is compatible, then for each  $(\theta_1, \ldots, \theta_k)$ , with  $\theta_h = 0$  if the h-th system is not compatible, a solution  $(\phi, \gamma)$  of (4) is obtained.

# 5.2. $P_d$ and $Q_d$ stochastic matrices.

Observing that the elements of  $P_d$  and  $Q_d$  satisfy (10), we can apply the procedure described in Section 5.1, with  $P=P_d$  and  $Q=Q_d$ , to check the compatibility of (12) and the coherence of  $(P_d,Q_d)$ . If (12) admits (at least) one strictly positive solution  $\{\bar{f}_i,\bar{g}_j,\ i\in I_0,\ j\in J_0\}$ , then  $(P_d,Q_d)$  is coherent. Moreover, being  $p_{ij}=0$ , if  $i\in I_0$ ,  $j\in J\setminus J_0$  and  $q_{ji}=0$ , if  $i\in I\setminus I_0$ ,  $j\in J_0$ , it is straightforward to verify that the values

(14) 
$$f_i = \bar{f_i}, i \in I_0, \ f_i = 0, i \in I \setminus I_0, \ g_j = \bar{g_j}, \ j \in J_0, \ g_j = 0, \ j \in J \setminus J_0,$$
 are a solution of system (4). Using (14), the assessment  $(P,Q)$  can be extended to a conditional probability  $\mathcal{P}$  on  $\mathcal{E} \times \mathcal{H}$ , where  $\mathcal{E}$  is the algebra generated by the family  $\mathcal{H} = \{\Omega, A_i, B_j, \ i \in I, \ j \in J\}$ , defining  $\mathcal{P}(A_i) = f_i, \mathcal{P}(B_j) = g_j$ . Recalling Corollary 4,  $\mathcal{P}$  is coherent if condition (2) is satisfied for the subscripts  $i, j$  such that  $f_i = g_j = 0$ , that is  $i \in I \setminus I_0, j \in J \setminus J_0$ . Since from Corollary 5 condition (2) on  $\mathcal{P}$  reduces to condition (3) applied to  $(P, Q)$ , then, for  $i \in I \setminus I_0$  and  $j \in J \setminus J_0$  condition (2) reduces to condition (3) applied to  $(P_a, Q_a)$ . To summarize,  $(P, Q)$  is coherent if (12) admits (at least) one strictly positive solution and  $(P_a, Q_a)$  satisfies condition (3).

## 5.3. $P_d$ and/or $Q_d$ not stochastic.

1 - Assume that  $P_d$  is not stochastic and denote by  $I_1 \subset I_0$  the set of indices i such that  $\sum_{j \in J_0} p_{ij} = 1$ . Then, for each  $i \in I_0 \setminus I_1$ , there exists  $j \in J \setminus J_0$  such that  $p_{ij} > 0$ , so that a necessary condition for compatibility of (4c) is  $f_i = 0$ ,  $i \in I_0 \setminus I_1$  and condition (15) below takes the place of (11).

(15) 
$$f_i = 0, \quad i \in I \setminus I_1; \quad g_j = 0, \quad j \in J \setminus J_0.$$

- 2 If  $I_1$  is empty, then  $f_i = 0$ ,  $i \in I$ , and system (4) does not have solutions; so that the assessment (P, Q) is not coherent.
- 3 If  $I_1$  is not empty, let  $P_1^d = \{p_{ij}\}$  and  $Q_1^d = \{q_{ji}\}$   $(i \in I_1, j \in J_0)$ . Observe that  $P_1^d$  is stochastic: then, if  $Q_1^d$  is stochastic too, we apply the procedure described in Section 5.2 to the partition of the matrices P and Q induced by the sets  $I_1$  and  $J_0$ .
- 4 If  $Q_1^d$  is not stochastic, denote by  $J_1 \subset J_0$  the set of indices such that  $\sum_{i \in I_1} q_{ji} = 1$ . Then, for each  $j \in J_0 \setminus J_1$ , there exists  $i \in I \setminus I_1$  such that  $q_{ji} > 0$ . So that compatibility of (4b) requires that  $g_j = 0$ ,  $j \in J_0 \setminus J_1$ , and the following condition takes the place of (15):

$$f_i = 0$$
,  $i \in I \setminus I_1$ ;  $g_j = 0$ ,  $j \in J \setminus J_1$ .

- 5 As in step 2, if  $J_1$  is empty, then system (4) is not compatible and (P, Q) is not coherent.
- 6- If  $J_1$  is not empty, let  $P_2^d = \{p_{ij}\}$  and  $Q_2^d = \{q_{ji}\}$   $(i \in I_1, j \in J_1)$ . Observe that  $Q_2^d$  is stochastic: then, if also  $P_2^d$  is stochastic we apply the procedure described in Section 5.2 to the partition of the matrices P and Q induced by the sets  $I_1$  and  $J_1$ .
- 7 Id  $P_2^d$  is not stochastic, then the previous steps 1-6 are applied to matrices  $P_2^d$  and  $Q_2^d$ .

The above algorithm stops at step 3 or step 6 if a pair of stochastic matrices, denoted by  $P^d$  and  $Q^d$ , is identified. The algorithm stops at step 2 or at step 5 if the condition  $f_i = 0$ ,  $i \in I$ , or  $g_j = 0$ ,  $j \in J$ , is respectively obtained.

In the first case the coherence of (P, Q) is checked by applying the procedure described in Section 5.2, with  $P_d = P^d$  and  $Q_d = Q^d$ . In the second case system (4) is not compatible and the assessment (P, Q) is not coherent.

#### REFERENCES

- [1] P. Berti E. Regazzini P. Rigo, de Finetti's coherence and complete predictive inferences, Working paper, IAMI 90.5, Milano, Italy, 1990.
- [2] P. Berti E. Regazzini P. Rigo, Coherent statistical inference and Bayes theorem, Annals of Statistics, 19 (1) (1991), pp. 366-381.
- [3] G. Coletti, Coherent numerical and ordinal probabilistic assessments, IEEE Trans. on Systems, Man and Cybernetics, 24 n.12 (1994), pp. 1747–1754.
- [4] G. Coletti R. Scozzafava, Characterization of coherent conditional probabilities, Techn. Rep. n.19, Univ. Perugia, 1994.
- [5] L. Crisma, *Prolungamenti di probabilità condizionate localmente coerenti*, Quad. n. 5/93, Dip. Mat. Appl. "B. de Finetti", Univ. Trieste, 1993.
- [6] Á. Császár, Sur la structure des espaces de probabilité conditionnelle, Acta Mathematica Academiae Scientiarum Hungaricae, 46 (1955), pp. 337–361.
- [7] B. de Finetti, Theory of probability, John Wiley, New York, 1974.
- [8] A. Gilio, Criterio di penalizzazione e condizioni di coerenza nella valutazione soggettiva della probabilità, Boll. Un. Mat. Ital., (7) 4-B (1990), pp. 645–660.
- [9] A. Gilio, Algorithms for precise and imprecise conditional probability assessments, Mathematical Models for Handling Partial Knowledge in Artificial Intelligence, eds. G. Coletti, D. Dubois, R. Scozzafava, Plenum Press, New York, 1995, pp. 231–254.
- [10] A. Gilio R. Scozzafava, Conditional events in probability assessment and revision, IEEE Trans. on Systems, Man and Cybernetics, 24 n. 12 (1994), pp. 1741–1746.
- [11] A. Gilio F. Spezzaferri, Knowledge integration for conditional probability assessments, Uncertainty in Artificial Intelligence, eds. D. Dubois, M.P. Wellman, B. D'Ambrosio and P. Smets, Morgan Kaufmann Publ., San Mateo, CA, 1992, pp. 98–103.
- [12] A.K. Gupta T. Varga, Characterization of joint density by conditional densities, Commun. Statist. Theory Meth., 19 (12) (1991), pp. 4643–4652.
- [13] S. Holzer, Sulla nozione di coerenza per le probabilità subordinate, Rendiconti dell'Istituto di Matematica dell'Università di Trieste, XVI (1984), pp. 46–62.
- [14] S. Holzer, On coherence and conditional prevision, Boll. Un. Mat. Ital., (6) 4 (1985), pp. 441–460.
- [15] R.S. Lehman, On confirmation and rational betting, Symbolic Logic, 20 (1955), pp. 251–262.
- [16] E. Regazzini, *Finitely additive conditional probabilities*, Rend. Sem. Mat. Fis. Milano, 55 (1985), pp. 69–89.

- [17] E. Regazzini, de Finetti's coherence and statistical inference, The Annals of Statistics, 15 (1987), pp. 845–864.
- [18] P. Rigo, Un teorema di estensione per probabilità condizionate finitamente additive, Atti della XXXIV Riunione Scientifica della S.I.S., Vol. 2 (2), Siena, 1988, pp. 27–34.
- [19] R. Scozzafava, Subjective probability versus belief functions in artificial intelligence, International Journal of General Systems, 22 (1994), pp. 197–206.
- [20] F. Spezzaferri, Reconstructing marginals from conditional distributions, Metron, XXXIX 3-4 (1981), pp. 161–176.
- [21] P.M. Williams, *Notes on conditional previsions*, Technical report, School of Mathematical and Physical Sciences, Univ. of Sussex, 1975.

Angelo Gilio, Dipartimento di Matematica, Università di Catania Viale A. Doria 6, 95125 Catania (ITALY)

Fulvio Spezzaferri, Dipartimento di Statistica, Università "La Sapienza", Piazzale A. Moro 5, 00185 Roma (ITALY)