Math-Net.Ru

Общероссийский математический портал

F. Mauch, On the existence of probability distributions with given marginals, *Теория вероятн. и ее примен.*, 2003, том 48, выпуск 3, 620–627

DOI: <https://doi.org/10.4213/tvp277>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.198.64.247

26 апреля 2023 г., 14:18:17

REFERENCES

1. Barron A. R. Entropy and the central limit theorem. — Ann. Probab., 1986, v. 14, № 1, p. 336–342.
2. Биллингсли П. Сходимость вероятностных мер. М.: Наука, 1977, 352 с.
3. Боровков А. А., Утес С. А. Об одном неравенстве и связанной с ним характеристизации нормального распределения. — Теория вероятн. и ее примен., 1983, т. 28, № 2, с. 209–218.
4. Brown L. D. A proof of the central limit theorem motivated by the Cramér–Rao inequality. — Statistics and Probability: Essays in Honor of C. R. Rao. Ed. by G. Kallianpur, P. R. Krishnaiah, and J. K. Ghosh. New York: North-Holland, 1982, p. 141–148.
5. Cacoullos Th. On upper and lower bounds for the variance of a function of a random variable. — Ann. Probab., 1982, v. 10, № 3, p. 799–809.
6. Chen L. H. Y. An inequality for the multivariate normal distribution. — J. Multivariate Anal., 1982, v. 12, № 2, p. 306–315.
7. Chen L. H. Y., Lou J. H. Asymptotic normality and convergence of eigenvalues. — Stochastic Process. Appl., 1990, v. 34, № 2, p. 197.
8. Chernoff H. A note on an inequality involving the normal distribution. — Ann. Probab., 1981, v. 9, № 3, p. 533–535.
9. Johnson O. T., Barron A. R. Fisher information inequalities and the central limit theorem. — math. PR/0111020.
10. Klaassen C. A. J. On an inequality of Chernoff. — Ann. Probab., 1985, v. 13, № 3, p. 966–974.
11. Nash J. Continuity of solutions of parabolic and elliptic equations. — Amer. J. Math., 1958, v. 80, p. 931–954.
12. Utev S. A. An application of integrodifferential inequalities in probability theory. — Siberian Adv. Math., 1992, v. 2, № 4, p. 164–199.

Поступила в редакцию
5.I.2001

Исправленный вариант
24.VI.2002

© 2003 г.

MAUCH F.*

ON THE EXISTENCE OF PROBABILITY DISTRIBUTIONS WITH GIVEN MARGINALS

Пусть $X = \{0, \dots, n-1\}$ и $\Gamma := \{(x_1, \dots, x_s) \in X^s : \sum_{\sigma=1}^s x_{\sigma} = n-1\}$. Дается простая характеристизация маргинальных распределений вероятностных законов на Γ , обладающих тем дополнительным свойством, что они образуют s -кортеж убывающих вероятностей на X . Эта характеристизация имеет интересное применение к асимптотическим спектрам в смысле Штрассена [5], [6]. Некоторые смежные вопросы обсуждаются в приложении.

Ключевые слова и фразы: вероятностный закон, маргинальные распределения, асимптотические спектры в смысле Штрассена, необходимые и достаточные условия.

* University of Constance, Germany.

1. Introduction and results. Let $X_1, \dots, X_s$ be nonempty finite sets and let Γ be a nonempty subset of $\prod X_\sigma$. For any probability distribution p on Γ we denote its marginal distributions on $X_1, \dots, X_s$ by $p^{(1)}, \dots, p^{(s)}$, respectively. We define

$$R := \{ \underline{p} = (p_1, \dots, p_s) : p_\sigma \text{ are probability distributions of } X_\sigma \},$$

$$M(\Gamma) := \{ (p^{(1)}, \dots, p^{(s)}) : p \text{ is a probability distribution on } \Gamma \}.$$

Obviously $M(\Gamma)$ and R are polytopes in $E := \prod \mathbf{R}^{X_\sigma}$ and $M(\Gamma) \subset R$. A necessary condition for $\underline{p} \in R$ to belong to $M(\Gamma)$ is for any $A_1 \subset X_1, \dots, A_s \subset X_s$

$$\sum p_\sigma(A_\sigma) \leq \max \left\{ \sum I_{A_\sigma}(x_\sigma) : (x_1, \dots, x_s) \in \Gamma \right\}, \quad (1)$$

where I_A denotes the indicator function of A . (Notice that the left-hand side equals $\int \sum I_{A_\sigma} p(d\underline{x})$, where the p_σ are the marginal distributions of p .) Dall'Aglio [2] and Kellerer [3] have shown that this condition is sufficient for $s = 2$, but not for $s \geq 3$. Nevertheless, by the duality theory for systems of linear inequalities one always has the necessary and sufficient condition that for any real functions f_σ on X_σ

$$\sum \int f_\sigma dp_\sigma \leq \max \left\{ \sum f_\sigma(x_\sigma) : (x_1, \dots, x_s) \in \Gamma \right\}. \quad (2)$$

(Again the necessity is trivial.) There also exists a fast algorithm for deciding this (the simplex algorithm). However, at least from a theoretical point of view, the above decision problem seems to be typically much harder for $s \geq 3$ than for $s = 2$.

In the main part of this paper we will treat the following parametric situation: Take a positive integer n and set $X := \{0, \dots, n-1\}$ and

$$\Gamma := \left\{ (x_1, \dots, x_s \in X^s : \sum_{\sigma=1}^s x_\sigma = n-1 \right\}. \quad (3)$$

Thus Γ is the simplex with vertices $(n-1, 0, \dots, 0), \dots, (0, \dots, 0, n-1)$ in the discrete hypercube X^s . Does $M := M(\Gamma)$ have a simple description?

If we take $f_\sigma(i) := i - (n-1)/s$ for all σ , then $\sum f_\sigma(x_\sigma) = 0$ on Γ . Hence by the (trivial) necessity of (2) we have $\sum \int f_\sigma dp_\sigma = 0$, i.e.,

$$h(\underline{p}) := \sum_\sigma \sum_i \left(i - \frac{n-1}{s} \right) p_\sigma(i) = 0 \quad \text{for any } \underline{p} \in M. \quad (4)$$

Call a probability distribution q on X decreasing, when $q(i+1) \leq q(i)$ for all i . Our main (and perhaps surprising) result is that for s -tuples of decreasing probability distributions condition (4) is already sufficient. Define

$$R^> := \{ \underline{p} \in R : p_\sigma \text{ is decreasing } \forall \sigma \}, \quad H := \{ \underline{p} \in (\mathbf{R}^n)^s : h(\underline{p}) = 0 \}.$$

$R^>$ is a polytope and H a hyperplane in $(\mathbf{R}^n)^s$. As we have seen above, $M \subset R \cap H$, hence $R^> \cap M \subset R^> \cap H$.

Theorem 1. $R^> \cap M = R^> \cap H$.

Thus for systems of decreasing probability distributions a single equation has to be tested to decide on marginality. We note that the restriction to decreasing probability distributions is necessary. In fact M is dramatically different from $R \cap H$ in the sense that $R \cap H$ has at most sn facets (since R is a product of s simplices and therefore has sn facets), whereas M has an exponential number of facets already for $s = 3$, as will be shown in the Appendix.

The phenomenon that the number of facets of $M(\Gamma)$ for certain Γ grows faster than any polynomial in n occurs even in the case $s = 2$. This is again shown in the Appendix and bases on Proposition 2 which states that if $\Gamma \subset X \times X$ is a partial order with a minimal and a maximal element then the number of concise facets of $M(\Gamma)$ equals the number of order ideals of X minus 2.

In this paper we will prove Theorem 1 for $s = 3$ and refer the reader to [4] for a proof of the general case.

An interesting application of Theorem 1 in the case $s = 3$ concerns asymptotic spectra in the sense of Strassen [5], [6]. These spectra are compact Hausdorff spaces, which control (in a certain asymptotic sense) degenerations as well as the computational complexity of finite dimensional bilinear maps. In [6] a singular 2-simplex ζ with values in Δ (the so-called support simplex) is constructed for certain asymptotic spectra Δ , and in [7] it is conjectured that the image of ζ coincides with Δ . At any rate the support simplex gives necessary conditions for the existence of degenerations between the bilinear maps under consideration.

Corollary 1. *Let k be an algebraically closed field and let f be the multiplication map of the algebra $k[T]/(T^m)$. Denote the standard 2-simplex by Θ and the asymptotic spectrum of f by $\Delta \subset \mathbb{R}$. Then the support simplex $\zeta: \Theta \rightarrow \Delta$ of f is given by*

$$\zeta(\theta) = \mu^{-(n-1)} \prod_{\sigma=1}^3 \left(\sum_{i=0}^{n-1} \mu^{i/\theta_\sigma} \right)^{\theta_\sigma}, \quad (5)$$

where μ is a unique positive solution of

$$\sum_{\sigma=1}^3 \frac{\sum_{i=0}^{n-1} i \mu^{i/\theta_\sigma}}{\sum_{i=0}^{n-1} \mu^{i/\theta_\sigma}} = n - 1. \quad (6)$$

By the Chinese Remainder Theorem any algebra $k[T]/(F)$ with $F \in k[T] \setminus \{0\}$ is isomorphic to a direct product of algebras of the type considered in Corollary 1. Therefore our corollary immediately gives the support simplex of the class of all algebras $k[T]/(F)$. (If the above mentioned conjecture is correct, we thus have computed the asymptotic spectrum for this class of bilinear maps. The asymptotic spectrum of a single algebra $k[T]/(F)$ has already been determined in [6].)

My thanks go to Volker Strassen for useful discussion about this matter.

2. Proof of Theorem 1. The inclusion $\langle \subset \rangle$ already being known we prove $\langle \supset \rangle$ by induction on n , the start $n = 1$ being clear. So assume $n > 1$.

It will suffice to show that the vertices of the polytope $R^> \cap H$ belong to $R^> \cap M$. Hence we determine the vertices of $R^> \cap H$. As H is a hyperplane, any such vertex is the intersection of an edge (a one-dimensional face) of $R^>$ with H . It is well known and easily seen that the vertices of the polytope of all decreasing probability distributions on $\{0, \dots, n-1\}$ are $r_1, \dots, r_n$, where

$$r_m(i) := \begin{cases} m^{-1} & \text{for } i < m, \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

Now $R^>$ is a direct product of three copies of this polytope. An edge of such a product is a product of one edge and two vertices in some order. In our case symmetry allows us to focus on edges of $R^>$ of the form $\{(1-\lambda)(r_{m_1}, r_{m_2}, r_{m_3}) + \lambda(r_{m_1}, r_{m_2}, r_{m'_3}): \lambda \in [0, 1]\}$ (where without loss of generality $m_3 < m'_3$). Hence we will consider vertices of $R^> \cap H$ of the form

$$\underline{p} = (1-\lambda)(r_{m_1}, r_{m_2}, r_{m_3}) + \lambda(r_{m_1}, r_{m_2}, r_{m'_3}), \quad m_3 < m'_3, \quad (8)$$

where λ is such that $\underline{p} \in H$. Now $h(r_{m_1}, r_{m_2}, r_{m_3}) = -(n-1) + \sum_{\sigma} \sum_i i r_{m_\sigma}(i) = -(n-1) + \sum_{\sigma} (m_\sigma - 1)/2$. Therefore the condition $\underline{p} \in H$ gives $0 = (1-\lambda) h(r_{m_1}, r_{m_2}, r_{m_3}) + \lambda h(r_{m_1}, r_{m_2}, r_{m'_3}) = -(n-1) + \frac{1}{2}((m_1-1) + (m_2-1) + (1-\lambda)(m_3-1) + \lambda(m'_3-1)) = -(n + \frac{1}{2}) + \frac{1}{2}(m_1 + m_2 + m_3 + \lambda(m'_3 - m_3))$, hence

$$\lambda = \frac{(2n+1) - m_1 - m_2 - m_3}{m'_3 - m_3}. \quad (9)$$

In order that $0 \leq \lambda \leq 1$ the following inequalities have to be satisfied:

$$m_1 + m_2 + m_3 \leq 2n + 1 \leq m_1 + m_2 + m'_3. \quad (10)$$

Now we take a vertex $\underline{p}$ of $R^> \cap H$. In order to apply the induction hypothesis we choose $k \in \{1, \dots, n-1\}$ and imbed the cube $\{0, \dots, n-k-1\}^3$ into the cube $\{0, \dots, n-1\}^3$ by means of the map $\varphi := \varphi_1 \times \varphi_2 \times \varphi_3$, where φ_1, φ_2 are the natural imbeddings of $\{0, \dots, n-k-1\}$ into $\{0, \dots, n-1\}$ and φ_3 is defined by $\varphi_3(i) := i+k$. This imbedding φ has the property that $\varphi^{-1}(\Gamma_n) = \Gamma_{n-k}$ (with obvious notation; similar notation will be used for R, M , and $R^>$).

Lemma 1. *Under the induction hypothesis the following is true: If $1 \leq k \leq n-1$ and $\underline{u} \in R \cap H$ is such that $u_1(n-1-i) = u_2(n-1-i) = u_3(i) = 0$ for $0 \leq i \leq k-1$ and u_1, u_2 are decreasing, whereas u_3 is decreasing on $\{k, \dots, n-1\}$, then $\underline{u} \in M$.*

P r o o f. Define $\varphi: R_{n-k} \rightarrow R_n$ by $\varphi(\underline{q}) := (\varphi_1(q_1), \varphi_2(q_2), \varphi_3(q_3))$. Then we have for any probability distribution q on Γ_{n-k} that the marginal triple of $\varphi(q)$ is the image under φ of the marginal triple of q . Hence φ maps M_{n-k} into M_n . Moreover

$$\begin{aligned} h_{n-k}(\underline{q}) &= -(n-k-1) + \sum_{\sigma} \sum_{i=0}^{n-k-1} i q_{\sigma}(i) = -(n-k-1) + \sum_{i=0}^{n-1} i(\varphi_1 q_1)(i) \\ &\quad + \sum_{i=0}^{n-1} i(\varphi_2 q_2)(i) + \sum_{i=0}^{n-1} (i-k)(\varphi_3 q_3)(i) = h_n(\varphi \underline{q}). \end{aligned}$$

Now let $\underline{u} \in R_n \cap H_n$ satisfy the assumptions of the lemma. It is clear that there exists a (unique) $\underline{q} \in R_{n-k}^>$ such that $\underline{u} = \varphi \underline{q}$. The above computation shows that $\underline{q}$ is even in $R_{n-k}^> \cap H_{n-k}$. By induction hypothesis $\underline{q} \in M_{n-k}$. Hence $\underline{u}$ is in M_n . Lemma 1 is proved.

It is clear that $\underline{p}$ does not satisfy the assumptions of the lemma, since p_3 is decreasing. Therefore we write $\underline{p}$ as a convex combination $\underline{p} = (1-\alpha)\underline{u} + \alpha\underline{v}$, where $\underline{u}, \underline{v} \in R \cap H$ are such that either both $\underline{u}$ and $\underline{v}$ satisfy the assumptions of the lemma (up to permutation of the three coordinates), or $\underline{u}$ does while $\underline{v}$ may be treated by the following lemma (or a permuted version of it).

Lemma 2. *Suppose $\underline{v} \in R$ is such that v_3 is the point mass at 0 (i.e., $v_3(0) = 1$). Then $\underline{v} \in M$ if and only if $v_2(i) = v_1(n-1-i)$ for all i .*

P r o o f. If v_3 is concentrated on 0, any probability distribution v on Γ with marginal triple $\underline{v}$ is supported by $\{(n-1-i, i, 0): 0 \leq i \leq n-1\}$. Hence $v_2(i) = v(n-1-i, i, 0) = v_1(n-1-i)$. Conversely, the probability distribution v on Γ defined by $v(n-1-i, i, 0) := v_1(n-1-i) = v_2(i)$ has marginal triple $\underline{v}$. Lemma 2 is proved.

We continue with the proof of the theorem by a case by case analysis. Remember that $\underline{p}$ is of the form (8) with (9) and (10).

C a s e 1: $p_1(n-1) > 0, p_2(n-1) > 0$. By (8) this means that $m_1 = m_2 = n$, hence by the first inequality of (10) we have $m_3 = 1$. From (9) we now get $\lambda = 0$, which in view of (8) says that p_3 is the point mass at 0. (The converse is also true: If p_3 is the point mass at 0, the assumption $m_3 < m'_3$ together with (8) gives $\lambda = 0$ and $m_3 = 1$, therefore by (9) one gets $m_1 = m_2 = n$.) Now we apply Lemma 2.

C a s e 2: $p_1(n-1) = p_2(n-1) = 0$. Here we use a decomposition $\underline{p} = (1-\alpha)\underline{u} + \alpha\underline{v}$, where $u_3(0) = 0$ and $v_3(0) = 1$ (i.e., v_3 is the point mass at 0). We will treat $\underline{u}$ by Lemma 1 and $\underline{v}$ by Lemma 2. Evaluating the equation $p_3 = (1-\alpha)u_3 + \alpha v_3$ at 0 we obtain $\alpha = p_3(0) > 0$ (p_3 is decreasing). Since $p_3(0) = 1$ is equivalent to Case 1 (as we have seen above), we conclude that $0 < \alpha < 1$. As a consequence $\underline{u}$ is determined by $\underline{v}$, i.e., by v_1 and v_2 . By Lemma 2 the condition $\underline{v} \in M$ means that $v_2(i) = v_1(n-1-i)$ for all i . Using this as a definition of v_2 in terms of v_1 it remains to define v_1 . The requirement that u_1 and u_2 be decreasing probability distributions actually dictates that v_1 is the rectangular distribution r_{n-m_2, m_1} , where

$$r_{l,m}(i) := \begin{cases} (m-l)^{-1} & \text{if } l \leq i < m, \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

for $0 \leq l < m \leq n$. (By (10) we have $n+1 \leq m_1 + m_2 + m'_3 - n \leq m_1 + m_2$, hence $n - m_2 < m_1$, hence r_{n-m_2, m_1} is well defined.) By the above definitions $\underline{v}$ is in M . Since $M \subset H$ this implies that $\underline{u}$ is in H . It remains to show that $\underline{u}$ is in R , i.e., that u_1 and u_2

are decreasing probability distributions. (Then we may apply Lemma 1 with $k = 1$ to $\underline{u}$ and finish Case 2.) By symmetry it suffices to treat u_1 . Since $p_1 = r_{m_1}$ and $v_1 = r_{n-m_2, m_1}$, we obtain

$$u_1 = \frac{r_{m_1} - \alpha r_{n-m_2, m_1}}{1 - \alpha}. \quad (12)$$

We only have to show that u_1 is nonnegative, i.e., that $\alpha \leq (m_1 + m_2 - n)/m_1$. (Then automatically u_1 is decreasing.) Now $\alpha = p_3(0)$. Hence by (8) and (9) we obtain

$$\begin{aligned} \alpha &= (1 - \lambda) r_{m_3}(0) + \lambda r_{m'_3}(0) = \frac{m'_3 - \lambda(m'_3 - m_3)}{m'_3 m_3} \\ &= \frac{1}{m'_3} \left(1 + \frac{m'_3 - (2n + 1) + m_1 + m_2}{m_3} \right). \end{aligned}$$

The numerator in the last expression is nonnegative by (10), hence by replacing m_3 by its minimal value 1 we obtain $\alpha \leq 1 + (-2n + m_1 + m_2)/m'_3$. Obviously, $-2n + m_1 + m_2 \leq 0$, hence by replacing m'_3 by its maximal value n we obtain $\alpha \leq (m_1 + m_2 - n)/n$. Since $n \geq m_1$ and α is positive, this implies the desired inequality and therefore finishes Case 2.

Before we treat the remaining two cases we argue that we may assume $0 < \lambda < 1$ (and hence strict inequalities in (10)). Indeed, when λ is either 0 or 1, the triple $\underline{p}$ consists of three rectangular distributions of the form (7). By symmetry we may assume that either $p_1(n-1) > 0$, $p_2(n-1) > 0$, or $p_1(n-1) = p_2(n-1) = 0$. These cases have already been handled in Case 1 and Case 2, respectively.

Without loss of generality we will also assume that $m_1 \leq m_2$.

Case 3: $p_1(n-1) = 0$, $p_2(n-1) > 0$, $p_3(n-1) = 0$. This will be treated in a similar way as Case 2, using a decomposition $\underline{p} = (1-\alpha)\underline{u} + \alpha\underline{v}$, where $u_2(0) = 0$ and $v_2(0) = 1$. We apply a permuted version of Lemma 1 to $\underline{u}$ and of Lemma 2 to $\underline{v}$. Evaluating the equation $p_2 = (1-\alpha)u_2 + \alpha v_2$ at 0 we obtain $\alpha = p_2(0) = 1/n$. As in Case 2 it suffices to define v_1 . We set $v_1 := r_{n-m'_3, m_1}$. By similar arguments as in Case 2 one proves that $r_{n-m'_3, m_1}$ is well defined and that it suffices to show $u_1(m_1 - 1) \geq 0$ and $u_3(m'_3 - 1) \geq 0$. This is equivalent to $n(m_1 + m'_3 - n) \geq m_1$ and $\lambda n(m_1 + m'_3 - n) \geq m'_3$, where λ is given by (9). The first inequality is clear, since $m_1 + m'_3 - n \geq 1$ by the well definedness of $r_{n-m'_3, m_1}$. Instead of the second inequality we will prove the stronger inequality $\lambda(m_1 + m'_3 - n) \geq 1$. In view of $m_2 = n$ this amounts to $(n - m_1 - m_3 + 1)(m_1 + m'_3 - n) \geq m'_3 - m_3$. Since $\lambda > 0$ the first factor on the left-hand side is ≥ 1 . We already know that the second factor is ≥ 1 . Therefore their product is greater or equal than their sum $m'_3 - m_3 + 1$. This finishes Case 3.

Case 4: $p_1(n-1) = 0$, $p_2(n-1) > 0$, $p_3(n-1) > 0$. Here we write $\underline{p} = (1-\alpha)\underline{u} + \alpha\underline{v}$ in such a way that Lemma 1 applies to $\underline{u}$ and a permuted version of Lemma 1 applies to $\underline{v}$. Explicitly we define

$$\begin{aligned} \alpha &:= \frac{m_1 + m_3 - 1}{n} && \text{(our assumption } \lambda > 0 \text{ implies } 0 < \alpha < 1), \\ u_1 &:= r_{m_1}, && v_1 := r_{m_1}, \\ u_2 &:= r_{n-m_1-m_3+1}, && v_2 := r_{n-m_1-m_3+1, n}, \\ u_3 &:= r_{m_3, n}, && v_3 := r_{m_3}. \end{aligned}$$

It is easily checked that $\underline{u}$ and $\underline{v}$ are well defined and lie in $R \cap H$. Now Lemma 1 applies to $\underline{u}$ with $k := m_3$ and to $\underline{v}$ with $k := n - m_1 - m_3 + 1$.

Since we have assumed $m_1 \leq m_2$, the four cases above exhaust all possibilities and the theorem is proved.

3. Appendix. First we are going to show that for $s = 3$ and Γ given by (3) the polytope $M = M(\Gamma)$ has at least $2^{n-2} - 1$ facets. Then we will consider the case $s = 2$ and construct subsets Γ of $X \times X$ such that $M(\Gamma)$ has exactly 2^{n-1} facets.

3.1. We begin with a general discussion of the faces of $M(\Gamma) \subset E = \prod \mathbf{R}^{X_\sigma}$ for arbitrary $\Gamma \subset \prod X_\sigma$.

First note that the set of vertices of $M(\Gamma)$ is

$$\{I_{\underline{x}} := (I_{x_1}, \dots, I_{x_s}) : \underline{x} \in \Gamma\},$$

where I_x for $x \in X_\sigma$ denotes the indicator of $\{x\}$. Indeed, $M(\Gamma)$ is obviously generated by this set and $(1 - I_{x_1}, \dots, 1 - I_{x_s}) \in E^*$ vanishes on $I_{\underline{x}}$, while it is positive on all other $I_{\underline{y}}$.

Any face S of $M(\Gamma)$ is generated by the vertices it contains, i.e., by $\{I_{\underline{x}}: \underline{x} \in \Gamma_S\}$, where Γ_S is a subset of Γ , uniquely determined by S . Since $M(\Gamma) \subset R$ lives in an affine subspace of E not containing 0, the face S is cut out by some $\underline{f} \in E^*$, so that $\underline{f} = 0$ on S , while $\underline{f} > 0$ on $M(\Gamma) \setminus S$. Then we have

$$\Gamma_S = \left\{ \underline{x}: \sum_{\sigma} f_{\sigma}(x_{\sigma}) = 0 \right\}.$$

If we call a subset of Γ closed, when it is of the form Γ_S for some face S of $M(\Gamma)$, we have a bijective correspondence $S \mapsto \Gamma_S$ between faces of $M(\Gamma)$ and closed subsets of Γ . Under this bijection the facets of $M(\Gamma)$ correspond to the maximal proper closed subsets of Γ .

Let π_{σ} denote the projection from Γ to X_{σ} , i.e., $\pi_{\sigma}(\underline{x}) = x_{\sigma}$ for $\underline{x} \in \Gamma$, and call Γ concise when all π_{σ} are surjective. We say that a face S of $M(\Gamma)$ is concise, when Γ_S is concise.

It is easy to see that a nonconcise facet is generated by $\pi_{\sigma}^{-1}(X \setminus \{x_{\sigma}\})$ for some $\sigma \in \{1, \dots, s\}$, $x_{\sigma} \in X_{\sigma}$. Thus the number of nonconcise facets is bounded by the sum of the sizes of the X_{σ} . Since our aim is to exhibit more than polynomial growth of the number of facets in certain situations, we will focus on concise facets.

3.2. Here we specialize to the case $s = 3$, $X = \{0, \dots, n-1\}$, $\Gamma := \{\underline{x} \in X^3: \sum_{\sigma} x_{\sigma} = n-1\}$. The projections π_{σ} are obviously surjective. We are going to prove that $M = M(\Gamma)$ has at least $2^{n-2} - 1$ concise facets.

For $\emptyset \neq A \subset \{0, \dots, n-3\}$ we define the subset $\bar{A}$ of Γ by

$$\begin{aligned} \bar{A} := & \{(i, n-3-i, 2): i \in A\} \cup \{(i, n-2-i, 1): 0 \leq i < n-2\} \\ & \cup \{(i, n-1-i, 0): i-1 \notin A\}. \end{aligned}$$

Thus, when Γ is considered as a discrete triangle, $\bar{A}$ is concentrated on the three lowest rows of Γ in the following way: In row 2 one has A , row 1 is filled completely, and row 0 carries the complement of the vertical copy of row 2.

Proposition 1. *There exists exactly one concise facet $S(A)$ of M such that $\bar{A} \subset \Gamma_{S(A)}$, and the map $A \mapsto S(A)$ is injective. Consequently, M has at least $2^{n-2} - 1$ concise facets.*

P r o o f. Suppose $\underline{f} \in E^*$ cuts out a concise facet S of M such that $\bar{A} \subset \Gamma_S$. In particular, for all $\underline{x} \in \bar{A}$ we have $\sum_{\sigma} f_{\sigma}(x_{\sigma}) = 0$ and for all $\underline{x} \in \Gamma$ we have $\sum_{\sigma} f_{\sigma}(x_{\sigma}) \geq 0$.

The conciseness of S means that Γ_S has surjective projections. Since $\bar{A} \subset \Gamma_S$, this is automatically true for π_1 and π_2 . Hence S is concise if and only if

$$f_3(k) = - \min_{i,j: i+j+k=n-1} f_1(i) + f_2(j) \quad \text{for } k \geq 3. \quad (13)$$

We may adjust $\underline{f}$ by adding a linear combination of $\underline{h}, (0, 1, -1), (1, -1, 0) \in E^*$ (where $\underline{h}$ has been defined in the introduction and 1 denotes the constant function 1 on X), since these linear forms vanish on M . Therefore we may assume that $f_3(0) = f_3(2)$, $f_3(1) = 0$ and $f_1(0) = 0$, and by scaling $f_3(0) \in \{\pm 1, 0\}$.

(row 1) For all $0 \leq i \leq n-2$ we have $(i, n-2-i, 1) \in \bar{A}$, therefore $f_1(i) + f_2(n-2-i) = 0$.

(row 2) For $1 \leq i \leq n-2$ we have $(i-1, n-2-i, 2) \in \Gamma$, hence from what we already know we get $0 \leq f_1(i-1) + f_2(n-2-i) + f_3(2) = f_1(i-1) - f_1(i) + f_3(0)$, i.e., $f_1(i) \leq f_1(i-1) + f_3(0)$.

Changing i to $i+1$ we obtain $f_1(i+1) \geq f_1(i) - f_3(0)$ for $0 \leq i \leq n-3$. Moreover, $f_1(i+1) = f_1(i) + f_3(0)$ when $i \in A$.

(row 0) For $0 \leq i \leq n-2$ we have $(i+1, n-2-i, 0) \in \Gamma$, hence $0 \leq f_1(i+1) + f_2(n-2-i) + f_3(0) = f_1(i+1) - f_1(i) + f_3(0)$, i.e., $f_1(i+1) \geq f_1(i) - f_3(0)$. Moreover, $f_1(i+1) = f_1(i) - f_3(0)$ when $i \notin A$.

Take $i \in A$. Then (row 2) and (row 0) give $f_3(0) = f_1(i+1) - f_1(i) \geq -f_3(0)$, hence $f_3 \geq 0$ and therefore $f_3(0) \in \{0, 1\}$.

Suppose $f_3(0) = 1$. Since $f_1(0) = 0$, (row 2) and (row 0) determine f_1 completely, (row 1) then fixes $f_2(i)$ for $i \leq n - 2$. But note that $(0, n - 1, 0) \in \bar{A}$, thus by what we already know we have $f_2(n - 1) = -1$. Hence f_2 and by (13) also f_3 are determined completely. It is clear (essentially again by (13)) that $\underline{f}$ cuts out a concise face of M .

A similar reasoning shows that $f_3(0) = 0$ leads to $\underline{f} = 0$, which is impossible.

Thus we have shown that there is exactly one concise face $S(A)$ of M such that $\bar{A} \subset \Gamma_s$. This face is necessarily a facet. Finally $S(A)$ together with the normalisation $f_3 = 1$ determines $\underline{f}$, which by (row 2) and (row 0) determines A .

3.3. Now we consider the case $s = 2$. We begin with an arbitrary $\Gamma \subset X_1 \times X_2$. Without loss of generality we may assume that $X_1 \cap X_2 = \emptyset$. We interpret $\Gamma \subset X_1 \times X_2$ as the set of edges of a bipartite graph on $X_1 \cup X_2$, which we denote by $G(\Gamma)$. Observe that Γ is automatically concise when $G(\Gamma)$ is connected.

Lemma 3. *Suppose $G(\Gamma)$ has k connected components. Then*

$$\dim M(\Gamma) = |X_1| + |X_2| - k - 1.$$

P r o o f. Since $0 \notin M(\Gamma)$, we have to prove

$$\dim \text{lin } M(\Gamma) = |X_1| + |X_2| - k. \tag{14}$$

Without loss of generality we may assume that Γ is concise. Since $G(\Gamma)$ has k connected components, we have decompositions $X_\sigma = \bigcup_{1 \leq j \leq k} X_{\sigma j}$ into nonempty subsets such that $\Gamma \subset \cup (X_{1j} \times X_{2j})$ and $(X_{1j} \cup X_{2j}, (X_{1j} \times X_{2j}) \cap \Gamma)$ is connected. Since both sides of (14) behave additively with respect to such a decomposition, it suffices to prove (14) for connected Γ , i.e., to show $\text{codim}(\text{lin } M(\Gamma)) = 1$ for connected Γ . Clearly, the linear form $(-1, 1)$ vanishes on $\text{lin } M(\Gamma)$. On the other hand, for any linear form (f_1, f_2) vanishing on $M(\Gamma)$ we have a decomposition

$$\Gamma = \bigcup_{\lambda \in \text{im } f_1} (f_1^{-1}(\lambda) \times f_2(f_1^{-1}(-\lambda))) \cap \Gamma.$$

(Indeed, for $(x_1, x_2) \in \Gamma$ let $f_1(x_1) = \lambda$. Then $f_2(x_2) = -\lambda$.) Since Γ is connected, this decomposition has only one term, so that f_1 and f_2 are constant and $f_2 = -f_1$.

Lemma 4. *Suppose that Γ is connected. Then the number of concise facets of $M(\Gamma)$ equals the number of pairs of nonempty subsets $A_1 \subset X_1$, $A_2 \subset X_2$ such that $(A_1 \times A_2) \cap \Gamma = \emptyset$ and $(A_1 \cup A_2^c, (A_1 \times A_2^c) \cap \Gamma)$ and $(A_1^c \cup A_2, (A_1^c \times A_2) \cap \Gamma)$ are connected.*

P r o o f. We construct a bijection between the two sets corresponding to the considered numbers.

First suppose S is a concise facet of $M(\Gamma)$. Then by Lemma 3 we have $\dim S = \dim M(\Gamma) - 1 = |X_1| + |X_2| - 3$, which implies (again by Lemma 3) that $G(\Gamma_S)$ has two connected components. Hence we have decompositions $X_\sigma = A_\sigma \cup A_\sigma^c$ such that $\Gamma_S \subset (A_1 \times A_2^c) \cup (A_1^c \times A_2)$ and $(A_1 \cup A_2^c, (A_1 \times A_2^c) \cap \Gamma)$ as well as $(A_1^c \cup A_2, (A_1^c \times A_2) \cap \Gamma)$ are connected and consist of at least one point.

If $A_1 = \emptyset$, then by the conciseness of S we have $X_2 = \pi_2(\Gamma_S) \subset \pi_2(X_1 \times A_2) = A_2$, hence $A_1 \cup A_2^c = \emptyset$. So $A_1 \neq \emptyset$. Analogously we get $A_2 \neq \emptyset$.

Now the linear form $(I_{A_1}, I_{A_2} - 1)$ vanishes on S . Since S is a facet the only linear forms vanishing on S are linear combinations of $(1, -1)$ and $(I_{A_1}, I_{A_2} - 1)$ and hence S is cut out by either $(I_{A_1}, I_{A_2} - 1)$ or $(-I_{A_1}, 1 - I_{A_2})$. Therefore either $(A_1^c \times A_2^c) \cap \Gamma = \emptyset$ or $(A_1 \times A_2) \cap \Gamma = \emptyset$. These two cases differ only in notation.

Conversely suppose A_1, A_2 are nonempty subsets of X_1, X_2 , respectively, such that $(A_1 \times A_2) \cap \Gamma = \emptyset$ and $(A_1 \times A_2^c) \cap \Gamma$ and $(A_1^c \times A_2) \cap \Gamma$ are connected. Since $(A_1 \times A_2) \cap \Gamma = \emptyset$ the linear form $(-I_{A_1}, 1 - I_{A_2})$ is ≥ 0 on $M(\Gamma)$ and cuts out $\Gamma_S := ((A_1 \times A_2^c) \cap \Gamma) \cup ((A_1^c \times A_2) \cap \Gamma)$ of Γ . Since $(A_1 \times A_2^c) \cap \Gamma$ and $(A_1^c \times A_2) \cap \Gamma$ are connected we have on the one hand from Lemma 3 that $\dim M(\Gamma_S) = \dim M(\Gamma) - 1$, therefore Γ_S is a facet of Γ . On the other hand we get $\pi_1(\Gamma_S) = A_1^c \cup A_1 = X_1$ and analogously $\pi_2(\Gamma_S) = X_2$. So Γ_S and the corresponding face S are concise facets.

It is easy to see that both defined maps are inverse to each other and give the bijection mentioned above.

Proposition 2. Suppose that $\Gamma \subset X \times X$ is a partial order with a minimal and a maximal elements. Then the number of concise facets of $M(\Gamma)$ coincides with the number of order ideals of X not equal to $\emptyset$ or X .

P r o o f. Since Γ has a minimal and a maximal elements, it is clear that Γ is connected. Now by Lemma 4 the number of concise facets equals the number of pairs A_1, A_2 of nonempty subsets of X such that $(A_1 \times A_2) \cap \Gamma = \emptyset$ and $(A_1 \cup A_2^c, (A_1 \times A_2^c) \cap \Gamma)$ and $(A_1^c \cup A_2, (A_1^c \times A_2) \cap \Gamma)$ are connected (in the following called property (A)).

Now it suffices to prove that the pair A_1, A_2 has property (A) if and only if A_2 is an order ideal not equal to $\emptyset$ or X and $A_1 = A_2^c$ (property (B)).

Suppose first that A_1, A_2 is a pair with property (A). $(A_1 \times A_2) \cap \Gamma = \emptyset$ implies that for each $x_1 \in A_1, x_2 \in A_2$ we have $x_1 \not\leq x_2$, in particular $A_1 \cap A_2 = \emptyset$. Next we assert $A_1 \cup A_2 = X$. Suppose $x \in A_1^c \cap A_2^c$. Since $(A_1 \cup A_2^c, (A_1 \times A_2^c) \cap \Gamma)$ is connected, there exists $x_1 \in A_1$ with $x_1 \leq x$. Since $(A_1^c \cup A_2, (A_1^c \times A_2) \cap \Gamma)$ is connected there exists $x_2 \in A_2$ with $x \leq x_2$. But now $x_1 \leq x_2$, hence $(x_1, x_2) \in (A_1 \times A_2) \cap \Gamma$ which is impossible. Hence $X = A_1 \cup A_2$. This implies immediately that A_2 is an order ideal: For each $x_2 \in A_2$ and $x_1 \leq x_2$ we have $x_1 \in A_1^c = A_2$. So A_1, A_2 has property (B).

Conversely, suppose A_1, A_2 is a pair with property (B). Since A_2 is nonempty the minimal element of X is in $A_2 = A_1^c$, therefore $(A_1^c \cup A_2, (A_1^c \times A_2) \cap \Gamma)$ is connected. Since $A_2 \neq X$ the maximal element of X is not in A_2 and therefore $(A_1 \cup A_2^c, (A_1 \times A_2^c) \cap \Gamma)$ is connected.

Applying Proposition 2 to $\{0, \dots, n-1\}$ with the natural order Γ_1 , we see that $M(\Gamma_1)$ has (only) $n-2$ concise facets.

But already for $X = \{0, \dots, n-1\}^2$ with the product order $\Gamma_2 = \Gamma_1 \times \Gamma_1$ we get $\binom{2n}{n}$ order ideals. (Represent an order ideal by a staircase, i.e., a weakly decreasing function from $\{0, \dots, n-1\}$ to $\{0, \dots, n\}$.) Hence $M(\Gamma_2) = M(\Gamma_1) \otimes M(\Gamma_1)$ has $\binom{2n}{n} - 2$ facets. In particular, we obtain the somewhat surprising result that a tensor product of two polytopes, each with a number of facets linear in the number of vertices, may have an almost exponential number of facets. For the number of order ideals of higher powers of $\{0, \dots, n-1\}$ consult [1].

A truly exponential number of facets is exhibited in the case of $X := \{0, \dots, n-1\}$ with the partial order Γ generated by $\{(0, 1), \dots, (0, n-2), (1, n-1), \dots, (n-2, n-1)\}$. Here we have 2^{n-2} order ideals apart from $\emptyset$ and X namely all $\{0\} \cup A$ with $A \subset \{1, \dots, n-2\}$.

If we define $X_1 := \{0, \dots, n-2\}$ and $X_2 := \{1, \dots, n-1\}$, then $M(\Gamma|_{X_1 \times X_2})$ has the same number of facets as Γ , but X_1, X_2 have only size $n-1$. This example is easily seen to give the maximum number of facets of $M(\Gamma)$ for any $\Gamma \subset Y_1 \times Y_2$ with $|Y_\sigma| \leq n-1$.

REFERENCES

1. Bolour A. Bounds on the number of integer valued monotone functions of k integer arguments. — Acta Arith., 1975, v. 28, p. 115–127.
2. Dall'Aglia G. Sulle distribuzioni doppie con margini assegnati sogette a delle limitazioni. — Giorn. Ist. Ital. Attuari, 1961, v. 24, p. 94–108.
3. Kellner H. Funktionen auf Produkträumen mit vorgegebenen Marginal-Funktionen. — Math. Ann., 1961, v. 144, p. 323–344.
4. Mauch F. Ein Randverteilungsproblem und seine Anwendung auf das asymptotische Spektrum bilinearer Abbildungen. — Doctoral Thesis. Allensbach: UFO, 1998.
5. Strassen V. The asymptotic spectrum of tensors. — J. Reine Angew. Math., 1988, v. 384, p. 102–152.
6. Strassen V. Degeneration and complexity of bilinear maps: Some asymptotic spectra. — J. Reine Angew. Math., 1991, v. 413, p. 127–180.
7. Strassen V. Algebra and complexity. — Progr. Math., 1994, v. 120, p. 429–446.

Поступила в редакцию
30.III.1999