Statistics and Probability Letters xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Termination of the iterative proportional fitting procedure

Q1 Fabian P. Reffel *

Institute of Mathematics, University of Augsburg, 86135 Augsburg, Germany

A R T I C L E I N F O

Article history:
Received 12 March 2014
Received in revised form 8 May 2014
Accepted 8 May 2014
Available online xxxx

MSC:
68W40
62H17

Keywords:
Iterative proportional fitting
IPF procedure
Alternating scaling
Termination

A B S T R A C T

The iterative proportional fitting procedure (IPF procedure) alternately fits a given nonnegative matrix to given positive row marginals and given positive column marginals. This paper proves that if the IPF procedure terminates, then this has to be within the first two steps.

1. Introduction

The IPF procedure aims to solve the following problem: Given a nonnegative $k \times \ell$ matrix A and positive marginals, find a nonnegative $k \times \ell$ matrix B which fulfills the given marginals and is biproportional to A. To this end, the IPF procedure generates a sequence of matrices $(A(t))$, called the IPF sequence, by alternately fitting rows and columns to match their respective marginals. The procedure is in use in many disciplines for problems such as calculating maximum likelihood estimators in graphical log-affine models (Lauritzen, 1996, Chapter 4.3.1), ranking webpages (Knight, 2008), determining passenger flows (McCord et al., 2010) or calculating seats in parliaments (Pukelsheim, 2014b, Chapter 14).

We say that the IPF procedure terminates in step T, when T is the smallest number such that the even-step IPF subsequence and the odd-step IPF subsequence replicate themselves after step T. That is, the two subsequences have converged in step T, respectively in step $T+1$. This paper shows that T only takes the values 0, 1, or 2.

Section 2 specifies the IPF procedure. The termination in the case of a converging IPF sequence is based on linear algebra and discussed in Section 3. This is the first main result. The second main result about the termination in case of a nonconverging IPF sequence relies on more sophisticated results and is presented in Section 4. Moreover, a characterization for the termination within two steps is given. Section 5 proposes perspectives for a further generalization.

In the sequel, all indices i belong to the set $\{1, \ldots, k\}$ whereas all indices j belong to the set $\{1, \ldots, \ell\}$. A subscript plus-sign indicates the summation over the index that would otherwise appear in its place. A set as a subscript denotes the summation over all entries belonging to that set, i.e. $r_I = \sum_{i \in I} r_i$. The transposed vector or transposed matrix is indicated by a prime. For all $n \in \mathbb{N}$, we define $\mathbb{1}_n := (1, \ldots, 1)' \in \mathbb{R}^n$.

2. IPF procedure

We specify the IPF procedure in full detail. The *IPF procedure* takes as input a nonnegative matrix $A \in \mathbb{R}^{k \times \ell}_{\geq 0}$ with positive row sums, $a_{i+} > 0$ for all i, and positive column sums, $a_{+j} > 0$ for all j, and two positive vectors $r \in \mathbb{R}^k_{>0}$ and $c \in \mathbb{R}^\ell_{>0}$.

* Tel.: +49 8215982202.
E-mail address: fabian.reffel@math.uni-augsburg.de.

http://dx.doi.org/10.1016/j.spl.2014.05.002
0167-7152/© 2014 Published by Elsevier B.V.

The matrix A is referred to as the *input matrix*, whereas the vector r is called the *row marginals* and the vector c is called the *column marginals*. The triple (A, r, c) forms the *input problem*.

The procedure is initialized by setting $A(0) := A$. Subsequently, the *IPF sequence* $(A(t))$ is calculated by iteratively repeating the following two steps:

- Odd steps $t + 1$ fit row sums to row marginals. To this end, all entries in the same row are multiplied by the same multiplier according to

$$a_{ij}(t+1) := \frac{r_i}{a_{i+}(t)} a_{ij}(t) \quad \text{for all entries } (i, j). \tag{1}$$

- Even steps $t + 2$ fit column sums to column marginals. To this end, all entries in the same column are multiplied by the same multiplier according to

$$a_{ij}(t+2) := \frac{c_j}{a_{+j}(t+1)} a_{ij}(t+1) \quad \text{for all entries } (i, j). \tag{2}$$

For all steps $t \geq 0$ and all entries (i, j) the inequality $a_{ij}(t) > 0$ holds if and only if $a_{ij} > 0$ holds. Consequently, all row sums $a_{i+}(t)$ and all column sums $a_{+j}(t)$ always stay positive. Thus, the IPF procedure is well defined. We say that the *IPF procedure converges* when the IPF sequence $(A(t))$ converges.

If $A(T) = A(T + 2)$ holds for some $T \in \mathbb{N}$, then $A(T + 1) = A(T + 3)$ holds as well. In this case, the even-step IPF subsequence $(A(t))$ and the odd-step IPF subsequence $(A(t + 1))$ stay constant for all even steps $t \geq T$. Therefore, the procedure can be terminated. We say that the IPF procedure *terminates* in step $T \in \mathbb{N}$ when T is the smallest natural number such that $A(T) = A(T + 2)$ holds.

As already mentioned by Rüschendorf (1995, p. 1164) and Vejnarová (2003, p. 585), the IPF procedure terminates after at most two steps if the input matrix A is of product form.

Example 2.1 (*Input Matrix of Product Form*)**.** Let (A, r, c) be an input problem such that for all entries (i, j) and some $u \in \mathbb{R}^k_{>0}$, $v \in \mathbb{R}^\ell_{>0}$ it holds $a_{ij} = u_i v_j$.

Then, the first three steps of the IPF procedure yield for all entries (i, j) the equations

$$a_{ij}(1) = \frac{r_i}{a_{i+}} a_{ij} = \frac{r_i}{\sum_q u_i v_q} u_i v_j = \frac{r_i}{v_+} v_j, \tag{3}$$

$$a_{ij}(2) = \frac{c_j}{a_{+j}(1)} a_{ij}(1) = \frac{c_j}{\sum_p \frac{r_p}{v_+} v_j} \frac{r_i}{v_+} v_j = \frac{r_i c_j}{r_+}, \tag{4}$$

$$a_{ij}(3) = \frac{r_i}{a_{i+}(2)} a_{ij}(2) = \frac{r_i}{\sum_q \frac{r_i c_q}{r_+}} \frac{r_i c_j}{r_+} = \frac{r_i c_j}{c_+}. \tag{5}$$

Further steps reproduce the matrices $A(2)$ and $A(3)$. Therefore, the IPF procedure terminates in step $T = 2$ at the latest. Moreover, if $c_+ = r_+$ holds, the IPF procedure converges within two steps. $\diamond$

3. Termination in case of convergent IPF procedure

In this section, we discuss the termination in case the IPF procedure is convergent. Then, the condition $A(T) = A(T + 2)$ is equivalent to $A(T) = A(T + 1)$. We show that T only attains the values 0, 1, or 2. To this end, we define $y^2 := (y_1^2, \ldots, y_\ell^2)$ for all $y \in \mathbb{R}^\ell$. The following lemma is crucial.

Lemma 3.1 (*Quadratic System of Equations*)**.** *Let $z \in \mathbb{R}^\ell_{>0}$ be given. Then $y \in \mathbb{R}^\ell$ is a solution of the system of equations*

$$z'y = z_+, \tag{6}$$

$$z'y^2 = z_+, \tag{7}$$

if and only if $y = \mathbb{1}_\ell$.

Proof. Let $y \in \mathbb{R}^\ell$ be a solution of Eq. (6). It holds

$$z'y^2 = \sum_j z_j y_j^2 = \sum_j z_j (2y_j - 1 + (y_j - 1)^2) = 2\sum_j z_j x_j - \sum_j z_j + \sum_j z_j (y_j - 1)^2 \tag{8}$$

$$= 2z'y - z_+ + \sum_j z_j (y_j - 1)^2 = z_+ + \sum_j z_j (y_j - 1)^2 \geq z_+. \tag{9}$$

If $y \neq \mathbb{1}_\ell$ holds, the above inequality is strict. However, this contradicts Eq. (7). Thus, $y = \mathbb{1}_\ell$ follows. $\square$