Limit points of the iterative scaling procedure

Erik Aas

Abstract

The iterative scaling procedure (ISP) is an algorithm which computes a sequence of matrices, starting from some given matrix. The objective is to find a matrix 'proportional' to the given matrix, having given row and column sums. In many cases, for example if the initial matrix is strictly positive, the sequence is convergent. It is known that the sequence has at most two limit points. When these are distinct, convergence to these two points can be slow. We give an efficient algorithm which finds the limit points, invoking the ISP only on subproblems for which the procedure is convergent.

1 Introduction

The *iterative scaling procedure* (ISP) is an algorithm which, given an $m \times n$ entrywise nonnegative matrix A and positive numbers $r_1, \ldots, r_m$, $c_1, \ldots, c_n$ attempts to find a matrix diagonally equivalent to A, having row sums r_i and column sums c_j. Two matrices A and A' are *diagonally equivalent* if there are sequences $(x_i^{(k)})$, $(y_j^{(k)})$, $k \geq 1$, $1 \leq i \leq m$, $1 \leq j \leq n$, of strictly positive numbers such that $a_{ij} = \lim_{k\to\infty} x_i^{(k)} a'_{ij} y_j^{(k)}$ for all i, j. (Of course, in some cases the limit can be omitted.) This is done by computing a sequence of $m \times n$ matrices by alternatingly scaling rows and columns, starting with A.

Depending on the initial data A, $\mathbf{r}$ and $\mathbf{c}$, this sequence might or might not converge. Csiszár and Tusnády [5] showed, in a much more general context, that the sequence has at most two limit points. Bregman [2] characterized completely the case when there is only one limit point. In this note we will describe the decomposition of the limit points in the general case.

The ISP has been applied in a variety of contexts, the most interesting of which perhaps being the ranking of webpages [6]. A discrete version of the algorithm is used by the Zürich City Council to distribute seats in parliamentary elections [1].

In Section 2, we define the ISP and state known convergence results. In Section 3, we give prove a characterization of the limit points, see Theorem 3. In Section 4, we summarize the results in form of an algorithm for computing the ISP limit points. In Section 5 we provide an example illustrating the main result. In the concluding Section 6, we state some questions related to the results proven here.

Acknowledgements
I am very thankful to Fabian Reffel who read an early version of this note, finding a serious error and providing many helpful suggestions. I also thank Kai-Friedrike Oelbermann and the Augsburg group headed by Prof. Friedrich Pukelsheim for telling me about the present problem and electoral methods in general. Svante Linusson and an anonymous referee provided many helpful suggestions.

2 Preliminaries

Throughout, A will denote a fixed nonnegative $m \times n$ matrix, and $\mathbf{r} = (r_1, \ldots, r_m), \mathbf{c} = (c_1, \ldots, c_n)$ fixed positive numbers. We further assume that there is no row or column in A containing only zeros, and that A is not the direct sum of two smaller matrices. For emphasis, we call $(A, \mathbf{r}, \mathbf{c})$ *positive* in this case. We will generally denote matrices by capital letters and their entries by the corresponding lower case letters; thus entry (i, j) in the matrix A is denoted a_{ij}. By the *row adjustment* (to $\mathbf{r}$) $\mathcal{R}(A)$ of A we mean the matrix whose (i, j) entry is $x_i a_{ij}$, where the *row multiplier* x_i is defined as $x_i = r_i / \sum_j a_{ij}$. We define the *column adjustment* $\mathcal{C}(A)$, and the *column multipliers* y_j (to $c_1, \ldots, c_n$) similarly. Define the *support* of a matrix M to be $S(M) := \{(i, j) : m_{ij} \neq 0\}$. By definition, $S(A) = S(\mathcal{R}(A)) = S(\mathcal{C}(A))$.

For example, if $\mathbf{r} = (7, 4)$, $\mathbf{c} = (3, 1)$, and $A = \begin{pmatrix} 2 & 5 \\ 4 & 0 \end{pmatrix}$, then $\mathcal{R}(A) = A$ and $\mathcal{C}(A) = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}$.

Note that $\mathcal{R}(A) = A = \mathcal{C}(A)$ in case A has both the desired row and column sums. The iterative scaling procedure consists of adjusting rows and columns alternatingly, starting with A. The iterates under the scaling procedure are defined to be

$$B^{(k+1)} := \mathcal{R}(C^{(k)}),$$

and

$$C^{(k)} = \mathcal{C}(B^{(k)})$$

for $k \geq 1$ and $B^{(1)} = \mathcal{R}(A)$.

We define $x_i^{(k)}$ as the row multipliers used when computing $B^{(k)}$, and $y_j^{(k)}$ as the column multipliers used when computing $C^{(k)}$.

Let us cite two theorems concerning the limits of $B^{(k)}$ and $C^{(k)}$.

Theorem 1. ([5], see also [4], section 5) For positive $A, \mathbf{r}, \mathbf{c}$, the sequences $B^{(k)}$ and $C^{(k)}$ are convergent.

Call these two limits, B and C respectively. We refer to these as the *ISP limits* (associated to $A, \mathbf{r}, \mathbf{c}$). Clearly $S(B) = S(C) \subseteq S(A), \mathcal{C}(B) = C$, and $\mathcal{R}(C) = B$.

Theorem 2. ([2]) We have $B = C$ if and only if there is some matrix M with row sums $\mathbf{r}$, column sums $\mathbf{c}$ and support a subset that of A; i.e. $S(M) \subseteq S(A)$.

The necessity part of Theorem 2 is easy to see. We will be concerned with the case when no such matrix M exists, i.e. when $B \neq C$.

We will use the following observation in what follows. If we scale the desired row sums by a common factor t, then this gives new ISP sequence $B'^{(k)}, C'^{(k)}$ closely related to the original sequence; we have $B'^{(k)} = tB^{(k)}$ and $C'^{(k)} = C^{(k)}$ for all $k \geq 1$. This is easily proved by induction on k.

Therefore we will not suppose that $\sum_i r_i = \sum_j c_j$. We can always reduce to the sums being equal by choosing t above as $\sum_j c_j / \sum_i r_i$, but it will be more natural not to do so.

In fact, if the limits are not equal, but $\sum_i r_i = \sum_j c_j$ and A is not the direct sum of two smaller matrices, then the support of the limit points is not equal to the support of the initial matrix, that is, some entries in the matrix tend to zero during the ISP. This is a consequence of Theorem 3 below.

For integer $n \geq 0$, we use the notation $[n] = \{1, \ldots, n\}$. Let $I \subseteq [m]$ be a set of row indices and $J \subseteq [n]$ a set of column indices. We call any such pair (I, J) a *block*.

A set $S = \{(I_k, J_k)\}_1^r$ of blocks is called a *splitting* if the sets I_k partition the set of rows, and the sets J_k partition the set of columns. The sets of rows and columns will always be those of the input matrix A in what follows.

An elementary refinement of a splitting consists of replacing a block (I_k, J_k) with (I', J') and (I'', J'') such that I', I'' partition I_k and J', J'' partition J_k. If the splitting S' is obtained by performing a sequence of elementary refinements on the splitting S then we say that S' is a *refinement* of S.

By the *decomposition* of a matrix B we will mean the splitting $I_1, \ldots, I_r$, $J_1, \ldots, J_r$ of the row and column sets of B such that $b_{ij} \neq 0$ implies $i \in I_k$ and $j \in J_k$ for some k, minimal with respect to refinement.

The algorithm to be described will find the decomposition of B (which of course is the same as that of C). The following observation from [7] shows that this suffices to compute B, running the ISP only on input for which it converges.

Proposition 1. Let A' be the matrix obtained from A by setting the (i, j) entry to 0 if $b_{ij} = 0$. Then the ISP limits of $(A, \mathbf{r}, \mathbf{c})$ and $(A', \mathbf{r}, \mathbf{c})$ coincide.

Let x_i and y_j be such that $x_i b_{ij} = c_{ij}$ and $c_{ij} y_j = b_{ij}$. Hence $x_i b_{ij} y_j = b_{ij}$ and thus $x_i y_j = 1$ whenever $(i, j) \in S(B)$. For subsets I of rows, define $r(I) := \sum_{i \in I} r_i$ and $c(J)$ similarly for subsets J of columns. Denote by $\mathcal{D}'$ the decomposition of B. It follows that for each block $(I, J) \in \mathcal{D}'$, x_i takes the same value for each $i \in I$ and similarly for y_j, $j \in J$. This common value is easily seen to be $x_i = 1/y_j = r(I)/c(J)$. The number $r(I)/c(J)$ will be called the *quotient* of the block (I, J). Therefore, $x_i^{(k)} \to r(I_p)/c(J_p)$ and $y_j^{(k)} \to c(J_p)/r(I_p)$ as $k \to \infty$ for $(i, j) \in I_p \times J_p$.

Thus, we conclude that, after reordering rows and columns suitably, the matrix B will look as follows.

$$
\begin{array}{c} \\ I_1 \\ I_2 \\ \vdots \\ I_r \end{array}
\begin{array}{c}
\begin{array}{cccc} J_1 & J_2 & \dots & J_r \end{array} \\
\begin{pmatrix}
B[I_1, J_1] & \bar{0} & \dots & \bar{0} \\
0 & B[I_2, J_2] & \dots & \bar{0} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \dots & B[I_r, J_r]
\end{pmatrix}
\end{array}
$$

We further assume the blocks are ordered so that $r(I_1)/c(J_1) \geq r(I_2)/c(J_2) \geq \dots$. Ties will be broken according to Lemma 6. A consequence of Theorem 3 will be that the barred zeros will in fact correspond to zeros in A.

Note that for each k, the submatrix $B[I_k, J_k]$ is a matrix with row sums r_i and column sums $(r(I_k)/c(J_k))c_j$. Similarly, $C[I_k, J_k]$ has column sums c_j and row sums $(c(J_k)/r(I_k))r_i$. Whenever $k \neq l$, we have $B[I_k, J_l] = C[I_k, J_l] = 0$.

We will say a block (I, J) is *feasible* if there is some $I \times J$ matrix M with row sums r_i, column sums $(r(I)/c(J))c_j$ and $S(M) \subseteq S(A)$. A splitting is feasible if all its blocks are feasible. So the decomposition of B is clearly feasible.

3 Proofs

Lemma 1. Let $p_1, \dots, p_n, q_1, \dots, q_n$ be positive real numbers. Then

$$\min_i \frac{p_i}{q_i} \leq \frac{p_1 + \cdots + p_n}{q_1 + \cdots + q_n} \leq \max_i \frac{p_i}{q_i}.$$

If any of the two inequalities is in fact an equality, then all the p_i/q_i are equal.

Proof. This follows by induction, the case $n = 2$ being easy. □

A more intuitive way to think about the fraction in the middle is as the average of the p_i's over the average of the q_i's. It is important to note that $\min_i p_i / \max_j q_j \leq (p_1 + \cdots + p_n)/(q_1 + \cdots + q_n)$ is a considerably weaker statement than Lemma 1.

We define the (bipartite) *graph* of A as follows. The vertices are the rows and columns of A, and row i and column j are adjacent in the graph if and only if $a_{ij} \neq 0$. If I is a set of rows, we denote by $N(I)$ the set of *neighbours* of I, i.e. the set $\{j : a_{ij} \neq 0 \text{ for some } i \in I\}$.

We will use the following simple generalization of a well-known theorem by Philip Hall (see eg. [3]). We formulate the lemma in terms of a submatrix of A for later convenience.

Lemma 2. Let (I', J') be a fixed block in A, and let $t = r(I')/c(J')$. Then there is a $I' \times J'$ matrix M with $S(M) \subseteq S(A)$ and row sums r_i, $i \in I'$, and column sums tc_j, $j \in J'$ if and only if there is no subset $I'' \subseteq I'$ of rows such that $r(I'') > tc(N(I'') \cap J')$.

For initial data $A, \mathbf{r}, \mathbf{c}$, we define $\varphi(A, \mathbf{r}, \mathbf{c})$ as the subset I of rows such that $\#I$ is maximal among those I maximizing $r(I)/c(N(I))$. Let us prove that $\varphi(A, \mathbf{r}, \mathbf{c})$ is well defined.

Lemma 3. If I_1 and I_2 satisfy the definition of $\varphi(A, \mathbf{r}, \mathbf{c})$, then $I_1 = I_2$.

Proof. Suppose $I_1 \neq I_2$. We have $r(I_1)+r(I_2) = r(I_1 \cup I_2)+r(I_1 \cap I_2)$ and $c(N(I_1)) + c(N(I_2)) = c(N(I_1) \cup N(I_2)) + c(N(I_1) \cap N(I_2)) \geq c(N(I_1 \cup I_2)) + c(N(I_1 \cap I_2))$.

Therefore

$$\frac{r(I_1) + r(I_2)}{c(N(I_1)) + c(N(I_2))} \leq \frac{r(I_1 \cup I_2) + r(I_1 \cap I_2)}{c(N(I_1 \cup I_2)) + c(N(I_1 \cap I_2))}.$$

By Lemma 1 either $I_1 \cup I_2$ or $I_1 \cap I_2$ shows that neither I_1 nor I_2 can satisfy the definition of $\varphi(A, r, c)$. □

Though it will follow from Theorem 3, it is interesting to note that we can prove directly that $(I, N(I))$ is feasible, where $I = \varphi(A, r, c)$.

Lemma 4. Let $I = \varphi(A, r, c)$. Then the block $(I, N(I))$ is feasible.

Proof. Suppose $(I, N(I))$ is not feasible. By Lemma 2 there is then some $I' \subseteq I$ such that $r(I') > (r(I)/c(N(I)))c(N(I'))$, or $r(I')/c(N(I')) > r(I)/c(N(I))$. But this contradicts the choice of I. □

Consider the decomposition $\mathcal{D}'$ of B and denote by $\Psi(B)$ the block obtained by merging all blocks with maximal quotient into a single block (which will have this same quotient). We will denote by $\mathcal{D}$ the splitting obtained from $\mathcal{D}'$ after this merge. Of course, the blocks in $\mathcal{D}$ are all feasible (since this is true for $\mathcal{D}'$).

Theorem 3. Suppose $A, \mathbf{r}, \mathbf{c}$ is positive, and let $I = \varphi(A, \mathbf{r}, \mathbf{c})$. Then $\Psi(B) = (I, N(I))$.

Proof. Let $(I_1, J_1) = \Psi(B)$. We wish to prove that $I_1 = I$ and $J_1 = N(I)$. We do this in two steps.

We clearly have $J_1 \subseteq N(I_1)$: otherwise $B[I_1, J_1]$ would have a zero column. For the same reason we have $J' \subseteq N(I')$ for each $(I', J') \in \mathcal{D}$.

We now prove that $N(I_1) \subseteq J_1$. Suppose this is not the case. Then there are $p \in I_1$, $q \notin J_1$ such that $a_{pq} \neq 0$. Denote by (I_2, J_2) the block in $\mathcal{D}$ such that $q \in J_2$.

We know that $x_p(k) \to r(I_1)/c(J_1)$ and $y_q(k) \to c(J_2)/r(I_2)$ as $k \to \infty$. Therefore

$$x_p(k)y_q(k) \to \frac{r(I_1)}{c(J_1)}\frac{c(J_2)}{r(I_2)} > 1.$$

Choose $\eta > 0$ and K such that $x_p(k)y_q(k) > 1 + \eta$ for all $k \geq K$. Hence $b_{pq}(K + n) > (1 + \eta)^n b_{pq}(K) \to \infty$ as $n \to \infty$. This contradicts the fact that all entries in the $B^{(k)}$ are bounded by $\max(r_1+\cdots+r_m, c_1+\cdots+c_n)$.

Thus $J_1 = N(I_1)$.

It follows from the previous step and the definition of I that $r(I)/c(N(I)) \geq r(I_1)/c(J_1)$.

Let $(I', J') \in \mathcal{D}$. We show that $r(I \cap I')/c(N(I \cap I') \cap J') \leq r(I')/c(J')$. Suppose not; then $N(I \cap I') \cap J' \neq J'$, since otherwise $r(I \cap I')/c(N(I \cap I') \cap J') = r(I \cap I')/c(J') \leq r(I')/c(J')$. Let $I'' = I \cap I'$. We have $r(I'') > (r(I')/c(J'))c(N(I'') \cap J')$, and thus by Lemma 2 the block (I', J') is not feasible, a contradiction.

Note that $r(I) = \sum r(I \cap I')$ and $c(N(I)) = \sum c(N(I) \cap J') \geq \sum c(N(I \cap I') \cap J')$, where the sums range over the blocks (I', J') of $\mathcal{D}$ such that $I \cap I' \neq \emptyset$.

We can therefore write

$$\frac{r(I)}{c(N(I))} = \frac{\sum r(I \cap I')}{\sum c(N(I) \cap J')} \leq \frac{\sum r(I \cap I')}{\sum c(N(I \cap I') \cap J')}. \tag{1}$$

We know from above and the definition of I that

$$\frac{r(I \cap I')}{c(N(I \cap I') \cap J')} \leq \frac{r(I')}{c(J')} \leq \frac{r(I_1)}{c(J_1)} = \frac{r(I_1)}{c(N(I_1))} \leq \frac{r(I)}{c(N(I))}$$

for each $(I', J') \in \mathcal{D}$ with $I \cap I' \neq \emptyset$, so by Lemma 1 all terms in each sum in the right hand side of (1) must be equal.

We now show that in fact there is only one term in (1), equal to $r(I \cap I_1)/c(N(I \cap I_1))$. Suppose to the contrary that we have $(I', J') \in \mathcal{D}$, $I' \cap I \neq \emptyset$, $(I', J') \neq (I_1, J_1)$, and $r(I \cap I')/c(N(I \cap I') \cap J') = r(I)/c(N(I))$.

Then we have $r(I \cap I')/c(N(I \cap I') \cap J') = r(I)/c(N(I)) \geq r(I_1)/c(J_1) > r(I')/c(J')$, and this contradicts (I', J') being feasible in the same manner as was done above. Thus the only term occurring is $r(I \cap I_1)/c(N(I \cap I_1) \cap J_1)$, and it equals $r(I)/c(N(I))$. Therefore $I \subseteq I_1$ and by the definition of I we thus conclude that $I = I_1$. □

Using Theorem 3, it is easy to find $\Psi(B)$, as we detail in Section 4.

The following lemma is a consequence of Theorem 2 in [8].

Lemma 5. If the only I satisfying $r(I) \geq c(N(I))$ are $I = \emptyset$ and $[m]$, then $S(A) = S(B)$.

The next lemma is needed for distinguishing blocks with the same multiplier (for example, the blocks in $\Psi(B)$). The proof is similar to that of Theorem 3, except we consider cumulative multipliers (to be defined) instead of incremental ones.

Lemma 6. Suppose $(A, \mathbf{r}, \mathbf{c})$ is positive, and let (I_k, J_k) be the blocks of the decomposition of B. Then the (I_k, J_k) can be ordered so that $N(I_k) \subseteq J_1 \cup \cdots \cup J_k$ for each k.

Proof. Note that a pair (i, j) is in the support of A but not of B if and only if the product of cumulative multipliers $X_i^{(k)} Y_j^{(k)} \to 0$ as $k \to \infty$. The cumulative multipliers are defined by $X_i^{(k)} = \prod_{l=1}^k x_i^{(l)}$ and $Y_j^{(k)} = \prod_{l=1}^k y_j^{(l)}$ (so that $a_{ij}^{(k)} = X_i^{(k)} a_{ij} Y_j^{(k)}$) for all k.

For blocks (I_a, J_a) and (I_b, J_b) we define $(I_a, J_a) \leq (I_b, J_b)$ if $S(A) \cap (I_b \times J_a) \neq \emptyset$. To prove the lemma it suffices to show that there is no cycle in $\leq$. To this end, suppose there is a cycle in $\leq$.

Then one can find a sequence of points $(i_1, j_1), (i_2, j_2), \ldots, (i_r, j_r)$ with the following properties

- r is even.
- For l even, $i_l = i_{l+1}$ and $j_l = j_{l-1}$, taking indices modulo r.
- If (i_l, j_l) does not belong to a block of B, then both its (cyclic) successor and predecessor do.

- $(i_r, j_r) \in S(A)\backslash S(B)$
- If l is odd, then $(i_l, j_l) \in S(B)$.

By assumption, for each k, $\prod_{l \text{ odd}} X_{i_l}^{(k)} Y_{j_l}^{(k)} = \prod_{l \text{ even}} X_{i_l}^{(k)} Y_{j_l}^{(k)}$. The limit of the left hand side is a product of b_{ij}/a_{ij} for $(i,j) \in S(B)$, hence strictly positive. But the limit of the right hand side is 0, since it contains at least one term $X_{i_l}^{(k)} Y_{j_l}^{(k)}$ with $(i_l, j_l) \in S(A)\backslash S(B)$, while all terms are clearly bounded. This contradiction finishes the proof. □

4 The algorithm

We now describe how to find the decomposition of B given A, $\mathbf{r}$, $\mathbf{c}$. This is done by repeatedly splitting blocks until this can be done no more, starting with the single block $([m], [n])$. For ease of notation we assume the current block under consideration is $([m], [n])$.

The splitting is naturally divided into two steps, I and II. In step I we look for a subblock of the current block with greatest quotient among all such subblocks. In step II we look for (proper) subblocks of the current block with same quotient as the current block. As observed earlier, we may assume the current block has quotient equal to 1.

Step I

First compute $I = \varphi(A, \mathbf{r}, \mathbf{c})$. This can be done in polynomial time in the size of A and the number of bits used to represent the input data, as follows. Let t be any positive number and consider the problem of finding a matrix with support contained in $S(A)$, row sums r_i and column sums $\leq t \cdot c_j$. For each fixed t this is a standard problem that can be solved with linear programming in polynomial time. Now, to compute $\varphi(A, \mathbf{r}, \mathbf{c})$, we want to find the largest t such that this problem has a solution with column sums *equal* to tc_j. If this property is satisfied for some t, it is of course also satisified for any smaller t. So one can perform binary search in t to find the rightmost point in the interval of t's such that the problem has such a solution. When we have found such a maximal t (which we can guarantee to be maximal by making sure it is in a sufficiently small interval), it is easy (for example, by successively forcing the values of variables in the linear program) to find $\varphi(A, \mathbf{r}, \mathbf{c})$ since it is the unique inclusion-wise maximal one among all solutions.

Now, if $r(I)/c(N(I)) > 1$, then split the current block into $(I, N(I))$ and $([m] - I, [n] - N(I))$. Otherwise, the current block is left unchanged.

When step I cannot be applied any further (to any block), we perform

Step II

As before, compute $I = \varphi(A, \mathbf{r}, \mathbf{c})$, forcing the solution to be different from $I = [m]$ if possible (as is easily done by modifying the previous linear program slightly). If this leads to a proper subset $I \subseteq [m]$ with $r(I)/c(N(I)) = 1$, then split the current block into $(I, N(I))$ and $([m] - I, [n] - N(I))$. Otherwise no such subset exists and the ISP limit of the current block will have the same support as A does in the current block, by Lemma 5.

Theorem 3 shows that after Step I, all blocks of B contained in some particular block of the ones found all have the same quotient. Lemma 6 show that when step II can not be applied further we will have found the decomposition of B.

Now, referring to Proposition 1, if we change A by setting the entries in $S(A)\backslash S(B)$ to 0, the ISP limit points will not change. However, numerical experiments suggest that convergence is much quicker than without the change. As a small example of this, take the matrix from the example below, set the entries outside the splitting found ($(1,3)$, $(1,4)$, $(2,1)$, $(2,3)$, and $(2,4)$) to zero. Then it takes about 2 iterations to get as close to (B,C) in the example, as $B(5)$ and $C(5)$ from earlier are from B and C.

5 An example

Let

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 7 & 2 \\ 1 & 1 & 9 & 6 \end{pmatrix}, \; r = (4,1,1,1), \text{ and } c = (1,2,2,2).$$

After 10 iterations we get the following matrices (rounded to 3 digits):

$$B(5) = \begin{pmatrix} 4.000 & 0.000 & 0.000 & 0.000 \\ \varepsilon_1 & 1.000 & 0.000 & 0.000 \\ \varepsilon_2 & 0.074 & 0.562 & 0.364 \\ \varepsilon_3 & 0.039 & 0.382 & 0.578 \end{pmatrix} \text{ and}$$

$$C(5) = \begin{pmatrix} 1.000 & 0.000 & 0.000 & 0.000 \\ \varepsilon_4 & 1.797 & 0.000 & 0.000 \\ \varepsilon_5 & 0.133 & 1.190 & 0.773 \\ \varepsilon_6 & 0.070 & 0.810 & 1.227 \end{pmatrix}.$$

Here, ε_1 and ε_4 are approximately 10^{-4}, and $\varepsilon_2, \varepsilon_3, \varepsilon_5, \varepsilon_6$ are approximately 10^{-6}.

The actual limit matrices are (rounded to 3 digits)

$$B = \begin{pmatrix} 4.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 1.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.604 & 0.396 \\ 0.000 & 0.000 & 0.396 & 0.604 \end{pmatrix} \text{ and } C = \begin{pmatrix} 1.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 2.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 1.209 & 0.791 \\ 0.000 & 0.000 & 0.791 & 1.209 \end{pmatrix}.$$

We have $\Psi(B) = (I_1, J_1)$ where $I_1 = \{1\}$, $J_1 = \{1\}$. One can check that we have $\varphi(A,r,c) = I_1$ and $J_1 = N(I_1)$ in this case, as expected.

In step I of the algorithm, we find the splitting consisting of the blocks (I_1, J_1) and (I_2, J_2) where $I_2 = \{2,3,4\}$ and $J_2 = \{2,3,4\}$.

In step II applied to the block (I_2, J_2) we find the subset $I_3 = \{2\} \subseteq I_2$ with the property $r(I_3) = 1 = (3/6)6 = (r(I_2)/c(J_2))c(N(I_2) \cap J_2)$. So the result of applying step II to (I_2, J_2) is (I_3, J_3) and $(I_4, J_4) := (I_2\backslash I_3, J_2\backslash J_3)$. Applying step II to any of the blocks (I_2, J_2), (I_3, J_3) or (I_4, J_4) yields nothing new, so the final splitting found is $\{(I_2, J_2), (I_3, J_3), (I_4, J_4)\}$, which coincides with the decomposition of B.

6 Future work

I would like to mention some related extensions and problems. First, most results above seem to carry over to the much more general setting of Theorem 5.2 in [4], but I have not explored this further.

A formalisation of the claim that convergence is quicker in the case $B = C$ would of course be very interesting.

There is a natural version of the ISP with continuous time, as follows. Let $\vartheta \in (0,1)$, and define $\mathcal{R}_\vartheta(A)_{ij} = x_i^\vartheta A_{ij}$, and $\mathcal{C}_\vartheta(A)_{ij} = A_{ij} y_j^\vartheta$, with x_i and y_j defined as before. Now we can define $F_{\alpha,\beta} = R_\alpha C_\beta$, and ask about the properties of $\lim_{\varepsilon \to 0} \lim_{n\to\infty} F^{(n)}_{\alpha\varepsilon,\beta\varepsilon}(A)$. The matrix $F_{\alpha,\beta}(A)$ will be diagonally equivalent to A and from this it follows (cf. [7]) that the limit $\lim_{\varepsilon \to 0} \lim_{n\to\infty} F^{(n)}_{\alpha\varepsilon,\beta\varepsilon}(A)$ will be the same as the ordinary ISP limit of A if the latter exists. It is not clear what happens in the general case when that limit does not exist.

Finally, the ISP can of course be defined for arbitrary, not necessarily nonnegative, A, r, c. Two issues arise in this case. One problem is that we may obtain matrices having marginals equal to 0 during the iteration. This could probably be avoided by using the continuous version described above; it seems reasonable such a system would repel from matrices having some marginal close to 0. For the discrete time version, the analogous convergence claims are not true; the initial data

$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $\mathbf{r} = (13, -12)$, $\mathbf{c} = (4, 6)$

gives a sequence with period 4. This cycle appears to be unstable in the input.

References

[1] Michel Balinski and Friedrich Pukelsheim (2006). Matrices and politics, *Festschrift for Tarmo Pukkila on his 60th Birthday* (Eds. E. Liski, J. Isotalo, S. Puntanen, G.P.H. Styan), 233-242.

[2] L. Bregman (1967). Proof of the convergence of Sheleikhovskii's method for a problem with transportation constraints. *USSR Computational Mathematics and Mathematical Physics*, 7:191-204.

[3] Richard Brualdi, Seymour Parter, Hans Schneider (1966). The diagonal equivalence of a nonnegative matrix to a stochastic matrix. *Journal of Mathematical Analysis and Applications*, 16, 31 - 50.

[4] Imre Csiszár and Paul Shields (2004). *Information theory and statistics*, now Publishers, Delft, section 5.2.

[5] Imre Csiszár and Gábor Tusnády, (1984). Information geometry and alternating minimization procedures. *Statistics and Decisions*, Supplement Issue, 1:205-237.

[6] Philip Knight (2008). The Sinkhorn-Knopp algorithm: convergence and applications, *SIAM Journal on Matrix Analysis and Applications*, 30:261-275.

[7] Oliver Pretzel (1980). Convergence of the iterative scaling procedure, *Journal of the London Mathematical Society*, 21, 379-384.

[8] Friedrich Pukelsheim (2012). An L^1-Analysis of the Iterative Proportional Fitting Procedure, preprint. Retrieved at `opus.bibliothek.uni-augsburg.de`.