Operations Research Letters 43 (2015) 74–79

Contents lists available at ScienceDirect

Operations Research Letters

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

Bounding stochastic dependence, joint mixability of matrices, and multidimensional bottleneck assignment problems

Utz-Uwe Haus

Institute for Operations Research, Department of Mathematics, ETH Zurich, Switzerland

ARTICLE INFO

Article history:
Received 24 July 2014
Received in revised form
24 November 2014
Accepted 29 November 2014
Available online 11 December 2014

Keywords:
Risk aggregation
VaR-bounds
Model uncertainty
Positive dependence
Bottleneck assignment

ABSTRACT

A matrix is jointly mixable if by permuting the entries in its columns all row sums can be made equal. If not jointly mixable we want to determine the smallest maximal and largest minimal row sum attainable. These values provide an approximation of the minimum variance problem for discrete distributions, estimating the α-quantile of an aggregate random variable with unknown dependence structure. We relate this NP-hard problem to the multidimensional bottleneck assignment problem and derive a PTAS in fixed dimension.

1. Introduction

The problem we are considering is the following: Given a matrix $A \in \mathbf{R}^{m\times d}$, we are interested in the best way of permuting entries in each column (independently) so that the maximal row sum is minimized, or so that the minimal row sum is maximized. Given d permutations $\Pi = (\pi_1, \dots, \pi_d) \in \mathfrak{S}(m)^d$ we denote by A^Π the matrix obtained from A by permuting column j by π_j, i.e. $A^\Pi_{\pi_j(i),j} = A_{i,j}$. The optimization problems are then

$$\gamma(A) := \min_{\Pi\in\mathfrak{S}(m)^d} \max_{1\leqslant i\leqslant m} \left\{ \sum_{j=1}^{d} A^\Pi_{i,j} \right\} \tag{1}$$

and

$$\beta(A) := \max_{\Pi\in\mathfrak{S}(m)^d} \min_{1\leqslant i\leqslant m} \left\{ \sum_{j=1}^{d} A^\Pi_{i,j} \right\}. \tag{2}$$

We note that aggregation operations other than $+$ are conceivable (e.g., min, max, $\times$), but will not be treated here.

This problem is motivated by an application in quantitative finance, but in fact arises whenever one needs to estimate the influence of stochastic dependence on a statistical problem: Consider an aggregate random variable L of the form $L = \sum_{i=1}^{d} L_i$, where the random variables L_i are possibly dependent. Denote by $F_L(x) = P(L \leqslant x)$ the distribution function of L. We are interested in computing the α-quantile (Value-at-Risk, VaR$_\alpha$) $F_L^{-1}(\alpha) = \inf\{x \in \mathbf{R} : F_L(x) \geqslant \alpha\}$, for $\alpha \in (0, 1)$. Often we have no data on the joint distribution L, but only on the marginal distributions F_j of the constituent random variables L_j, and we also lack information on the dependence structure between them.

In the following we will assume that the marginal distributions are discrete, or have been approximated from below and from above as described in [10]: For F_j the generalized inverse is $F_j^{-1}(\alpha) = \sup\{x \in \mathbf{R} : F_j(x) \leqslant \alpha\}$. Consider a discretization in $N+1$ points. Compute the values $q_r^j = F_j^{-1}(r/N)$ for $r \in \{0, 1, \dots, N\}$. Denoting by $1_{[a,b)}$ the characteristic function on the interval $[a, b)$,

$$\underline{F_j}(x) = \frac{1}{N}\sum_{r=0}^{N-1} 1_{[q_r^j,+\infty)}(x) \quad\text{and}\quad \overline{F_j}(x) = \frac{1}{N}\sum_{r=1}^{N} 1_{[q_r^j,+\infty)}(x),$$

provide discrete approximations of F_j with $\underline{F_j} \geqslant F_j \geqslant \overline{F_j}$.

Dependence among the individual F_j will manifest itself in the way the values $q_r^j = F_j^{-1}(r/N)$ are appearing in the matrix

$$A = \begin{pmatrix} q_0^1 & \cdots & q_0^d \\ \vdots & & \vdots \\ q_N^1 & \cdots & q_N^d \end{pmatrix}.$$

In particular, the row sums may vary significantly: Consider $d = 2$ and the uniform discrete distribution on $\{0, \dots, N\}$. If L_1 and L_2 are comonotonic (i.e. there is perfect positive dependence

E-mail address: uhaus@ifor.math.ethz.ch.

http://dx.doi.org/10.1016/j.orl.2014.11.009
0167-6377/

among the random variables), then $(q_0^1, \dots, q_N^1) = (q_0^2, \dots, q_N^2)$ with row sums $\{0, 2, \dots, 2N\}$. If, on the other hand, F_1 and F_2 are countermonotonic (perfect negative dependence among the random variables), then $(q_0^1, \dots, q_N^1) = (q_N^2, \dots, q_0^2)$, and all row sums are equal to N. If we want to find an upper bound for $F_L^{-1}(\alpha)$ we need to consider matrices with entries q_r^j for $\frac{r}{N} \geqslant \alpha$, and for lower bounds matrices constructed from q_r^j for $\frac{r}{N} \leqslant \alpha$ and each time minimize the variance of the row sums of A. This intuition is made exact by a representation theorem of Rüschendorf [14, Theorem 2], showing that for discrete distribution functions, and due to the uniform discretization inherent in our definition of $\underline{F_j}$ and $\overline{F_j}$, solving the minimum variance problem amounts to determining $\gamma(A) - \beta(A)$ for the matrix A, since it is enough to minimize over the set of all rearrangements of the F_j. We refer to [10,11,15,9] for recent applications and to [14,12,1] for more details on the general concept of rearrangements of functions.

Besides computing (or approximating) $\gamma(A)$ and $\beta(A)$, one is also interested in deciding whether for a given matrix $\gamma(A) = \beta(A)$. We will call such a matrix *jointly mixable*, in analogy with the definition of this concept by Wang and Wang [15] for distribution functions.

In this paper we show that deciding joint mixability is a strongly $\mathcal{NP}$-complete problem, even for a fixed number of columns, but can be solved using dynamic programming in pseudopolynomial time for a fixed number of rows. We show that the algorithm proposed by Puccetti et al. in [10] to compute $\gamma(A)$ and $\beta(A)$ is not an exact method unless $\mathcal{NP} \subseteq \mathcal{ZPP}$, despite its impressive computational success [2]. Finally, for matrices in fixed (column) dimension we present a polynomial-time approximation scheme.

2. An application

A typical application is the following (see [2]): Under the Basel II and III regulatory framework for banking supervision, large international banks are allowed to come up with internal models for the calculation of risk capital. For operational risk the so-called Loss Distribution Approach gives them full freedom concerning the stochastic modeling assumptions used. The resulting risk capital must correspond to a 99.9% quantile of the aggregated loss data over a year. This corresponds to computing the Value-at-Risk $\mathrm{VaR}_{0.999}(L)$ at $\alpha = 0.999$ for an aggregate loss random variable $L = \sum_{i=1}^d L_i$, but makes no requirements on the interdependence between the individual loss random variables L_i corresponding to the individual business lines: Assumptions made in the calculation must only be plausible and well founded. Estimating the upper bound and lower bound of the VaR over all possible dependence structures is hence relevant both from the regulator's point of view, as well as from the bank's point of view, to estimate worst case hidden risks in the models presented under the Loss Distribution Approach.

3. Complexity

It is known that for two columns the joint mixability problem is solvable explicitly (see references in [13]). This is also apparent by recognizing that the computation of $\gamma(A)$ can be understood as solving a multidimensional bottleneck assignment problem. The multidimensional bottleneck assignment problem asks for the computation of

$$\min_{\pi_1,\dots,\pi_d} \max_{1\leqslant i\leqslant m} c_{\pi_1(i),\dots,\pi_d(i)}$$

for a $\underbrace{m \times \cdots \times m}_{d}$ cost table C. Defining $c_{i_1,\dots,i_d} = A_{i_1,1} + \cdots + A_{i_d,d}$ we see that $\gamma(A)$ can be computed by solving a multidimensional bottleneck assignment problem. Using Observation 1 below we can similarly compute $\beta(A)$ and thus check joint mixability.

In dimension 2, the bottleneck assignment problem models the following problem: Given a set of workers and a set of tasks, where the time of worker i performing task j is c_{ij}, find a simultaneous assignment of all workers to all tasks such that the maximal time spent by any worker (the bottleneck of the schedule) is minimized. Fulkerson et al. showed that the 2-dimensional bottleneck assignment problem can be transformed into a linear assignment problem, and thus is polynomially solvable [3].

The multi-dimensional bottleneck assignment problem of assigning (equal-sized) crews of workers to (equal-sized) groups of tasks is much harder; even some versions of the 3-dimensional bottleneck assignment problem with extra constraints on the matrix A do not admit a polynomial time approximation scheme [7].

By adding $\mu = -\min_{1\leqslant i\leqslant m,\ 1\leqslant j\leqslant d} A_{ij}$ to each entry of A we can always shift the matrix to make the smallest entry equal to zero, changing all row sums by $+\mu \cdot d$. For convenience we will hence restrict our attention to integral, nonnegative matrices. Assuming that integrality is not a major restriction, since rational matrices can without loss of generality be scaled to become integral, and rational matrices provide a dense subset of the real matrices that could arise in discretizing distribution functions.

First note that β and γ are related as follows:

Observation 1. *Let $A \in \mathbf{Z}^{m\times d}$, and denote by*

$$l := \max_{1\leqslant i\leqslant m, 1\leqslant j\leqslant d} A_{ij}$$

its largest entry. Define A' by $A'_{ij} = l - A_{ij}$. Then $\beta(A) = d \cdot l - \gamma(A')$.

Hence we only ever need to consider one of the two values.

Let us consider some matrix A^{Π} with row sums $r_1^{\Pi}, \dots, r_m^{\Pi}$. By summing up $\sum_{i=1}^m r_i^{\Pi} = \sum_{i=1}^m \sum_{j=1}^d A_{ij}$, so if $r_1^{\Pi} = \cdots = r_m^{\Pi}$ the candidate value is $s = \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^d A_{ij}$. It always holds that $\max_{1\leqslant i\leqslant m} r_i^{\Pi} \geqslant s$, so also $\gamma(A) \geqslant s$. But if $\gamma(A) = s$ then all row sums need to be equal since s is the average row sum, so A is jointly mixable. This observation shows that deciding joint mixability of A and computing β or γ are polynomially equivalent:

Observation 2. *Let $A \in \mathbf{Z}^{m\times d}$. A is jointly mixable if and only if $\gamma(A) = \beta(A) = \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^d A_{ij}$.*

It turns out that this is sufficient for showing linear time decidability of joint mixability if the entries of A are restricted to at most two values: Those can be mapped to $\{0, 1\}$, and then the algorithm used in the proof below provides a linear time check for joint mixability:

Theorem 1. *Let $A \in \{0, 1\}^{m\times d}$. A is jointly mixable if and only if $m \mid \sum_{1\leqslant i\leqslant m,\ 1\leqslant j\leqslant d} A_{ij}$. The permutation achieving the joint mix can be computed in linear time $\mathcal{O}(m \cdot d)$.*

Proof. "⇒" Let $s = \sum_{i=1}^m \sum_{j=1}^d A_{ij}$. If $m \nmid s$ then by Observation 2 the matrix A cannot be jointly mixable.

"⇐" Assume $m \mid \sum_{1\leqslant i\leqslant m, 1\leqslant j\leqslant d} A_{ij}$. We need to permute the columns of A such that exactly $r = \frac{s}{m} \in \{0, \dots, d\}$ entries in each row have value 1.

This can always be done: Define for $i \in \{1, \dots, m\}$ the *defect* $\delta(i) = r - \sum_{j=1}^d A_{ij}$ and $\phi = \sum_{i=1}^m |\delta(i)|$ the total defect. Clearly, $\phi = 0$ if and only if all row sums of the matrix are equal to r.

Starting with $j = 2$ define $S_j = \{i \in \{1, \dots, m\} : \delta(i) > 0, A_{ij} = 1\}$ and $D_j = \{i \in \{1, \dots, m\} : \delta(i) < 0, A_{ij} = 0\}$. If $S_j \neq \varnothing$ and $D_j \neq \varnothing$ let $t_j = \min\{|S_j|, |D_j|\}$ and swap the entries of column $A_{\cdot j}$ indexed by the largest t_j entries of S_j with those indexed by the smallest t_j entries of D_j. Repeat in increasing order, for all $j \leqslant d$.