Matrix Completion for the Independence Model

Kaie Kubjas
joint work with Zvi Rosen
arXiv:1407.3254

Aalto University

October 13, 2014
Main problem

Problem

*Given some entries of a matrix, is it possible to add the missing entries so that the matrix has rank 1, its entries sum to one, and it is nonnegative?*
Example

For example, the partial matrix

\[
\begin{pmatrix}
0.16 &  \\
0.09 & 0.04 \\
0.01 &
\end{pmatrix}
\]

has a unique completion:

\[
\begin{pmatrix}
0.16 & 0.12 & 0.08 & 0.04 \\
0.12 & 0.09 & 0.06 & 0.03 \\
0.08 & 0.06 & 0.04 & 0.02 \\
0.04 & 0.03 & 0.02 & 0.01
\end{pmatrix}
\]
For example, the partial matrix

\[
\begin{pmatrix}
0.16 & 0.09 \\
0.09 & 0.04 \\
0.04 & 0.01 \\
\end{pmatrix}
\]

has a unique completion:

\[
\begin{pmatrix}
0.16 & 0.12 & 0.08 & 0.04 \\
0.12 & 0.09 & 0.06 & 0.03 \\
0.08 & 0.06 & 0.04 & 0.02 \\
0.04 & 0.03 & 0.02 & 0.01 \\
\end{pmatrix}
\]

On the other hand, perturbing any entry of the original matrix by \( \epsilon > 0 \) makes the matrix have no eligible completions, and perturbing any entry by \( \epsilon < 0 \) introduces an infinite number of completions.
Let $X$ and $Y$ be two independent discrete random variables with $m$ and $n$ states respectively, i.e.

$$Pr(X = i, Y = j) = Pr(X = i) \cdot Pr(Y = j)$$

for all $i, j$. 
Let $X$ and $Y$ be two independent discrete random variables with $m$ and $n$ states respectively, i.e.

$$Pr(X = i, Y = j) = Pr(X = i) \cdot Pr(Y = j)$$

for all $i, j$. Their joint probabilities are recorded in the matrix

$$P = \begin{pmatrix}
Pr(X = 1) \\
Pr(X = 2) \\
\vdots \\
Pr(X = m)
\end{pmatrix}
\begin{pmatrix}
Pr(Y = 1) & Pr(Y = 2) & \cdots & Pr(Y = n)
\end{pmatrix},$$

which is rank 1, nonnegative, and its entries sum to one.
Motivation

Suppose that probabilities $Pr(X = i, Y = j)$ are measurable only for certain pairs $(i, j)$. A situation in which this might arise in applications is a pair of compounds in a laboratory that only react when in certain states. A complete answer to our question will allow us to reject a hypothesis of independence of the events $X$ and $Y$, based only on this collection of probabilities.
Suppose that probabilities \( Pr(X = i, Y = j) \) are measurable only for certain pairs \((i, j)\).

A situation in which this might arise in applications is a pair of compounds in a laboratory that only react when in certain states.
Suppose that probabilities \( Pr(X = i, Y = j) \) are measurable only for certain pairs \((i, j)\).

A situation in which this might arise in applications is a pair of compounds in a laboratory that only react when in certain states.

A complete answer to our question will allow us to reject a hypothesis of independence of the events \( X \) and \( Y \), based only on this collection of probabilities.
Example

Let $M$ be the partial probability matrix given by:

$$M = \begin{pmatrix} a & \ \ \\
& b \end{pmatrix}.$$
Example

Let $M$ be the partial probability matrix given by:

$$M = \begin{pmatrix} a \\ b \end{pmatrix}.$$

Set the off-diagonal entries to $x$ and $ab/x$ and set the sum of all entries $a + ab/x + x + b$ equal to 1. The equivalent quadratic equation is $x^2 + (a + b - 1)x + ab = 0$. 
Example

Let $M$ be the partial probability matrix given by:

$$M = \begin{pmatrix} a \\ b \end{pmatrix}.$$ 

Set the off-diagonal entries to $x$ and $ab/x$ and set the sum of all entries $a + ab/x + x + b$ equal to 1. The equivalent quadratic equation is $x^2 + (a + b - 1)x + ab = 0$. In order for a real solution for $x$ to exist, the discriminant must be $\geq 0$, i.e.

$$(a + b - 1)^2 - 4ab \geq 0.$$ 

This inequality, along with the requirement that $a + b \leq 1$ and both $a, b > 0$, is sufficient to guarantee that $x$ gives a completion of $M$. 
Figure: The colored region corresponds to completable probability matrices of $2 \times 2$ matrices with diagonal entries $a, b$ observed.
Theorem (K., Rosen)

Let $M$ be an $n \times n$ partial probability matrix, where $n \geq 2$, with nonnegative observed entries along the diagonal:

$$M = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}.$$

Then $M$ is completable if and only if $\sum_{i=1}^{n} \sqrt{a_i} \leq 1$, or equivalently, $\|(a_1, \ldots, a_n)\|_{1/2} \leq 1$. In the special case $\sum_{i=1}^{n} \sqrt{a_i} = 1$, the partial matrix $M$ has a unique completion.
Diagonal partial matrices

Corollary

Let $\sum_{i=1}^{n} \sqrt{a_i} < 1$. For $n = 2$, the probability matrix $M$ has exactly two completions. If $n > 2$, then the set of completions of $M$ is $(n - 2)$-dimensional.
Corollary

Let \( \sum_{i=1}^{n} \sqrt{a_i} < 1 \). For \( n = 2 \), the probability matrix \( M \) has exactly two completions. If \( n > 2 \), then the set of completions of \( M \) is \((n - 2)\)-dimensional.

Figure: Each curve represents values of \( u \) that parametrize a completion of the \( 3 \times 3 \) diagonal partial matrix with \( 1/9, 1/10, 1/16, 1/36, 1/64, \) and \( 1/150 \) on the diagonal.
The analysis of the previous theorem works to derive the constraint for the $2 \times 2$ diagonal probability matrix as in the example. Assuming $a, b \geq 0$:
\[
\sqrt{a} + \sqrt{b} \leq 1 \iff a + b + 2\sqrt{ab} \leq 1 \iff 2\sqrt{ab} \leq 1 - a - b
\]
\[
\iff 4ab \leq (1 - a - b)^2 \text{ and } 0 \leq 1 - a - b.
\]
The analysis of the previous theorem works to derive the constraint for the $2 \times 2$ diagonal probability matrix as in the example. Assuming $a, b \geq 0$:

\[
\sqrt{a} + \sqrt{b} \leq 1 \iff a + b + 2\sqrt{ab} \leq 1 \iff 2\sqrt{ab} \leq 1 - a - b \\
\iff 4ab \leq (1 - a - b)^2 \text{ and } 0 \leq 1 - a - b.
\]
The analysis of the previous theorem works to derive the constraint for the $2 \times 2$ diagonal probability matrix as in the example. Assuming $a, b \geq 0$:

$$\sqrt{a} + \sqrt{b} \leq 1 \iff a + b + 2\sqrt{ab} \leq 1 \iff 2\sqrt{ab} \leq 1 - a - b$$

$$\iff 4ab \leq (1 - a - b)^2 \text{ and } 0 \leq 1 - a - b.$$
Example: General partial matrices

\[
\begin{pmatrix}
\ast & \ast & \ast \\
.06 & .09 & \ast \\
.08 & \ast & \ast \\
\ast & \ast & .15
\end{pmatrix}
\]
Example: General partial matrices

\[
\begin{pmatrix}
  * & * & * \\
  .06 & .09 & * \\
  .08 & * & *
\end{pmatrix}
\rightarrow
\begin{pmatrix}
  .08 & * & * \\
  * & * & .15
\end{pmatrix}
\]
Example: General partial matrices

\[
\begin{pmatrix}
* & * & * \\
.06 & .09 & * \\
.08 & * & * \\
* & * & .15
\end{pmatrix}
\rightarrow
\begin{pmatrix}
.06 & .09 & * \\
.08 & * & * \\
* & * & .15
\end{pmatrix}
\rightarrow
\begin{pmatrix}
.06 & .09 & * \\
.08 & .12 & * \\
* & * & .15
\end{pmatrix}
\]

\[
\sqrt{.35} + \sqrt{.15} = .98 < 1
\]
Example: General partial matrices

\[
\begin{pmatrix}
\ast & \ast & \ast \\
.06 & .09 & \ast \\
.08 & \ast & \ast \\
\ast & \ast & .15 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
\ast & \ast & \ast \\
.06 & .09 & \ast \\
.08 & \ast & \ast \\
\ast & \ast & .15 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
\ast & \ast & \ast \\
.06 & .09 & \ast \\
.08 & .12 & \ast \\
\ast & \ast & .15 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
\ast & \ast & \ast \\
.35 & \ast \\
\ast & .15 \\
\end{pmatrix}
\]
Example: General partial matrices

\[
\begin{pmatrix}
* & * & * \\
.06 & .09 & * \\
.08 & * & * \\
* & * & .15
\end{pmatrix} 
\rightarrow 
\begin{pmatrix}
* & * & .15 \\
* & * & .15
\end{pmatrix}
\rightarrow 
\begin{pmatrix}
.06 & .09 & * \\
.08 & .12 & * \\
* & * & .15
\end{pmatrix} 
\rightarrow 
\begin{pmatrix}
.35 & * \\
* & .15
\end{pmatrix}
\]

\[
\sqrt{.35} + \sqrt{.15} = .98 < 1
\]
Theorem

Let $M$ be a feasible partial matrix such that after carefully removing zeros, it has $s$ blocks. Let $b_i$ be the sum of the entries in the $i$-th block after completing by $2 \times 2$ minors. If $s = 1$, then $M$ is completable if and only if $b_1 = 1$. For $s > 1$, the partial matrix $M$ is completable to a probability matrix if and only if:

$$\sum_{i=1}^{s} \sqrt{b_i} \leq 1.$$
Example: General partial matrices

The probability matrix

\[
\begin{pmatrix}
  x_{11} & x_{12} \\
  x_{21} & x_{33}
\end{pmatrix}
\]

with all observed entries nonnegative has a completion if and only if

\[
\sqrt{x_{11} + x_{12} + x_{21} + x_{12}x_{21}/x_{11}} + \sqrt{x_{33}} \leq 1.
\]
Suppose we are given a partial probability tensor $T \in (\mathbb{R}^n)^\otimes d$ with nonnegative observed entries $a_i$ along the diagonal, i.e. we have $t_{ii...i} = a_i$ for $1 \leq i \leq n$, and all other entries unobserved. Then $T$ is completable if and only if

$$\sum_{i=1}^{n} a_i^{1/d} \leq 1.$$
Semialgebraic sets

Proposition

Partial tensors of fixed type which can be completed to rank-1 probability tensors form a semialgebraic set.
Partial tensors of fixed type which can be completed to rank-1 probability tensors form a semialgebraic set.

There exists a unique irreducible polynomial $f$ of degree $d^{n-1}$ with constant term 1 that vanishes on the boundary of diagonal partial tensors which can be completed to rank-1 probability tensors. The semialgebraic description takes the form $f \geq 0$, coordinates $\geq 0$ plus additional inequalities that separate our set from other chambers in the region defined by $f \geq 0$.
The region defined by

$$\sqrt{a} + \sqrt{b} + \sqrt{c} \leq 1$$

is the same as

$$((1 - S_1)^2 - 4S_2)^2 - 64S_3 \geq 0$$

together with

$$a, b, c \geq 0,$$
$$1 - S_1 \geq 0,$$
$$(1 - S_1)^2 - 4S_2 \geq 0,$$

where $S_1 = a + b + c$, $S_2 = ab + bc + ca$ and $S_3 = abc$. 
$3 \times 3$ diagonal matrices
**Figure:** The colored region corresponds to completable probability matrices of $3 \times 3$ matrices with diagonal entries observed.
Completion algorithm for two blocks

Consider the partial matrix

\[
\begin{pmatrix}
.06 & .09 \\
.08 & .15
\end{pmatrix}
\]

Step 1: Add in the missing entries using $2 \times 2$ minors:

\[
\begin{pmatrix}
.06 & .09 \\
.08 & .12
\end{pmatrix}
\]\n
Step 2: Add in $X$ and the rest of entries:

\[
\begin{pmatrix}
.06 & .09 & X \\
.08 & .12 & X \\
X & X & X
\end{pmatrix}
\]
Completion algorithm for two blocks

Consider the partial matrix

\[
\begin{pmatrix}
\cdot.06 & \cdot.09 \\
\cdot.08 & \cdot.15
\end{pmatrix}
\]

**Step 1:** Add in the missing entries using $2 \times 2$ minors:

\[
\begin{pmatrix}
\cdot.06 & \cdot.08 \\
\cdot.09 & \cdot.12 \\
\cdot.09 & \cdot.15
\end{pmatrix}
\]
Completion algorithm for two blocks

Consider the partial matrix

\[
\begin{pmatrix}
.06 & .09 \\
.08 \\
.09 & .12 \\
.15 \\
\end{pmatrix}
\]

**Step 1:** Add in the missing entries using $2 \times 2$ minors:

\[
\begin{pmatrix}
.06 & .08 \\
.09 & .12 \\
.15 \\
\end{pmatrix}
\]

**Step 2:** Add in $X$ and the rest of entries:

\[
\begin{pmatrix}
.06 & .08 & X \\
.09 & .12 & 1.5X \\
.009/X & .012/X & .15 \\
\end{pmatrix}
\]
Completion algorithm for two blocks

**Step 3:** Set $\sum p_{ij} = 1$ and solve for $X$:

$$(.06 + .08 + .09 + .12 + .15) + X + 1.5X + .009/X + .012/X = 1$$

$\Rightarrow .5 + 2.5X + .021/X = 1 \Rightarrow 2.5X^2 - .5X + .021 = 0$$
Completion algorithm for two blocks

**Step 3:** Set $\sum p_{ij} = 1$ and solve for $X$:

$$(.06 + .08 + .09 + .12 + .15) + X + 1.5X + .009/X + .012/X = 1$$

$\Rightarrow .5 + 2.5X + .021/X = 1 \Rightarrow 2.5X^2 - .5X + .021 = 0$

The two solutions for $X$ yield the following two completions:

$$\begin{pmatrix}
.06 & .08 & .06 \\
.09 & .12 & .09 \\
.15 & .2 & .15
\end{pmatrix} \quad \begin{pmatrix}
.06 & .08 & .14 \\
.09 & .12 & .21 \\
.06 & .09 & .15
\end{pmatrix}$$
Proposition

Let \( A = \text{diag}(a_1, \ldots, a_n) \), such that \( n > 2 \) and \( S = \sum \sqrt{a_i} < 1 \). Then, a completion of the matrix is given by:

\[
\mathbf{u} = \left( \frac{\sqrt{a_1}}{S} + t, \frac{\sqrt{a_2}}{S} - t, \frac{\sqrt{a_3}}{S}, \ldots, \frac{\sqrt{a_n}}{S} \right),
\]

where \( t \) is one of the solutions to the following quadratic equation:

\[
\left( \frac{\sqrt{a_1} + \sqrt{a_2}}{S} \right) t^2 + \left( a_2 - a_1 - \left( \frac{\sqrt{a_1} + \sqrt{a_2}}{S} \right)^2 \right) t
\]
\[
+ \left( \frac{a_1 \sqrt{a_2} + a_2 \sqrt{a_1}}{S} - \frac{\sqrt{a_1 a_2} (\sqrt{a_1} + \sqrt{a_2})}{S^3} \right) = 0,
\]

both of which lie in the interval \([-\sqrt{a_1}/S, \sqrt{a_2}/S]\).
Completion algorithm for more blocks

Example

We want to find a completion of $A = \text{diag}(1/4, 1/25, 1/36)$ that minimizes the Pearson $\chi^2$ distance from the uniform distribution:

$$d = \frac{1}{n^2} \sum_{i,j} \left( p_{ij} - \frac{1}{n^2} \right)^2 = \frac{1}{n^2} \sum_{i,j} \left( u_i v_j - \frac{1}{n^2} \right)^2.$$

This can be done using Lagrange multipliers. The minimum is achieved at

$$M = \begin{pmatrix} 0.250 & 0.049 & 0.215 \\ 0.204 & 0.040 & 0.176 \\ 0.032 & 0.006 & 0.028 \end{pmatrix} \quad \text{and} \quad M^T.$$

The Pearson $\chi^2$ distance from the uniform distribution is 0.683.
2 × 2 × 2 tensors

1. (Size 1) Any singleton, e.g. \( p_{000} \). The only condition is \( p_{000} \leq 1 \).

2. (Size 2) Three orbits of pairs:
   1. \( p_{000}, p_{001} \): \( p_{000} + p_{001} \leq 1 \).
   2. \( p_{000}, p_{011} \): \( \sqrt{p_{000}} + \sqrt{p_{011}} \leq 1 \).
   3. \( p_{000}, p_{111} \): \( 3\sqrt{p_{000}} + 3\sqrt{p_{111}} \leq 1 \).

3. (Size 3) Three orbits of triples:
   1. \( p_{000}, p_{001}, p_{010} \): \( p_{000} + p_{001} + p_{010} + (p_{001}p_{010}/p_{000}) \leq 1 \).
   2. \( p_{000}, p_{001}, p_{110} \): \( \sqrt{p_{000}} + \sqrt{p_{001}} + \sqrt{p_{110} + p_{001}p_{110}/p_{000}} \leq 1 \).
   3. \( p_{000}, p_{101}, p_{011} \): The tensor is completable if and only if the equation
      \[
      x^3 + (p_{000} + p_{101} + p_{011} - 1)x^2 + 
      (p_{000}p_{101} + p_{000}p_{011} + p_{101}p_{011})x + p_{000}p_{101}p_{011} = 0
      \]
      has a root in the interval \([0, 1]\).
Thank you!