Notes on Bayesian updating from financial statements

If we have more parameters to be estimated than data, we often say the problem is under-identified. However, this is a common problem in accounting. To wit, we often ask what activities did the organization engage in based on our reading of their financial statements. We know there is a simple linear relation between the recognized accounts and transactions

$$Ay = x$$

where A is an $m \times n$ matrix of ± 1 and 0 representing simple journal entries in its columns and adjustments to individual accounts via the journal entries in its rows, y is the transaction amount vector, and x is the change in the account balance vector over the period of interest (Arya, et al [2000]). Since there are only $m-1$ linearly independent rows (due to the balancing property of accounting) and m (the number of accounts) is almost surely less than n (the number of transactions we seek to estimate) we're unable to invert from x to recover y. Do we give up? If so, we might be forced to conclude financial statements fail even this simplest of tests.

Rather, we might take a page from physicists (Jaynes [2003]) and allow our prior knowledge to assist estimation of y. Of course, this is what decision theory also recommends. If our prior or background knowledge provides a sense of the first two moments for y, then the Gaussian or normal distribution is our maximum entropy prior. Maximum entropy implies that we fully utilize our background knowledge but don't use background knowledge we don't have (Jaynes [2003], ch. 11). That is, maximum entropy priors combined with Bayesian revision make efficient usage of both background knowledge and information from the data (in this case, the financial statements). As in previously discussed accounting examples, background knowledge reflects potential equilibria based on strategic interaction of various, relevant economic agents and accounting recognition choices for summarizing these interactions.

Suppose our background knowledge $\Im$ is completely summarized by

$$E\left[y \mid \Im\right] = \mu$$

and

$$Var\left[y \mid \Im\right] = \Sigma$$

then our maximum entropy prior distribution is

$$p\left(y \mid \Im\right) \sim N\left(\mu, \Sigma\right)$$

and the posterior distribution for transactions, y, conditional on the financial statements, x, is

$$\begin{aligned} & p\left(y \mid x, \Im\right) \\ \sim \quad & N\left(\mu + \Sigma A_0^T \left(A_0 \Sigma A_0^T\right)^{-1} A_0 \left(y^p - \mu\right), \Sigma - \Sigma A_0^T \left(A_0 \Sigma A_0^T\right)^{-1} A_0 \Sigma\right) \end{aligned}$$

where $N(\cdot)$ refers to the Gaussian or normal distribution with mean vector denoted by the first term, and variance-covariance matrix denoted by the second term, A_0 is A after dropping one row and y^p is any consistent solution to $Ay = x$ (for example, form any spanning tree from a directed graph of $Ay = x$ and solve for y^p). For the special case where $\Sigma = \sigma^2 I$ (perhaps unlikely but nonetheless illuminating), this simplifies to

$$p(y \mid x, \Im) \sim N\left(P_{R(A)} y^p + \left(I - P_{R(A)}\right)\mu, \sigma^2\left(I - P_{R(A)}\right)\right)$$

where $P_{R(A)} = A_0^T\left(A_0 A_0^T\right)^{-1} A_0$ (projection into the rowspace of A), and then $I - P_{R(A)}$ is the projection into the nullspace of A.[1]

The logic behind the belief updating above is as follows:

- The relation between transactions y and changes in account balances x is

  $$Ay = x$$

  or for the reduced form (to eliminate redundancy in the accounts) where A_0 drops a row from A and x_0 drops the corresponding element from x

  $$A_0 y = x_0$$

- Since changes in account balances x_0 are a linear combination of y, if the vector of transactions y is normally distributed then the vector x_0 is also normally distributed.

  $$y \sim N(\mu, \Sigma)$$

  implies[2]

  $$x_0 \sim N(E[x_0], Var[x_0])$$

  or

  $$x_0 \sim N\left(A_0\mu, A_0\Sigma A_0^T\right)$$

  The latter is derived as follows. Since $x_0 = A_0 y$

  $$\begin{aligned} E[x_0] &= E[A_0 y] \\ &= A_0 E[y] = A_0\mu \end{aligned}$$

[1] In the general case, we could work with the subspaces (and projections) of $A_0\Gamma$ where $\Sigma = \Gamma\Gamma^T$ (the Cholesky decomposition of Σ) and the transformed data $z \equiv \Gamma^{-1}y \sim N\left(\Gamma^{-1}\mu, I\right)$ (Arya, Fellingham, and Schroeder [2000]). Then, the posterior distribution of z conditional on the financial statements x is

$$p(z \mid x, \Im) \sim N\left(P_{R(A_0\Gamma)} z^p + \left(I - P_{R(A_0\Gamma)}\right)\mu_z, I - P_{R(A_0\Gamma)}\right)$$

where $z^p = \Gamma^{-1}y^p$ and $\mu_z = \Gamma^{-1}\mu$. From this we can recover the above posterior distribution of y conditional on x via the inverse transformation $y = \Gamma z$.

[2] Dropping a row from A ensures the existence of $\left(A_0\Sigma A_0^T\right)^{-1}$ when we update beliefs.

and

$$\begin{aligned}
Var\left[x_0\right] &= Var\left[A_0 y\right] \\
&= E\left[\left(A_0 y - A_0 \mu\right)\left(A_0 y - A_0 \mu\right)^T\right] \\
&= E\left[A_0\left(y-\mu\right)\left(y-\mu\right)^T A_0^T\right] \\
&= A_0 E\left[\left(y-\mu\right)\left(y-\mu\right)^T\right] A_0^T \\
&= A_0 Var\left[y\right] A_0^T \\
&= A_0 \Sigma A_0^T
\end{aligned}$$

$$\begin{aligned}
Cov\left[y, x_0\right] &= Cov\left[y, A_0 y\right] \\
&= E\left[\left(y-\mu\right)\left(A_0 y - A_0 \mu\right)^T\right] \\
&= E\left[\left(y-\mu\right)\left(y-\mu\right)^T A_0^T\right] \\
&= E\left[\left(y-\mu\right)\left(y-\mu\right)^T\right] A_0^T \\
&= Var\left[y\right] A_0^T \\
&= \Sigma A_0^T
\end{aligned}$$

$$\begin{aligned}
Cov\left[x_0, y\right] &= Cov\left[A_0 y, y\right] \\
&= E\left[\left(A_0 y - A_0 \mu\right)\left(y-\mu\right)^T\right] \\
&= E\left[A_0\left(y-\mu\right)\left(y-\mu\right)^T\right] \\
&= A_0 E\left[\left(y-\mu\right)\left(y-\mu\right)^T\right] \\
&= A_0 Var\left[y\right] \\
&= A_0 \Sigma
\end{aligned}$$

Hence, the joint distribution for y and x_0 is

$$\left[\begin{array}{c} y \\ x_0 \end{array}\right] \sim N\left(\left[\begin{array}{c} E\left[y\right] \\ E\left[x_0\right] \end{array}\right], \left[\begin{array}{cc} Var\left[y\right] & Cov\left[y, x_0\right] \\ Cov\left[x_0, y\right] & Var\left[x\right] \end{array}\right]\right)$$

or

$$\left[\begin{array}{c} y \\ x_0 \end{array}\right] \sim N\left(\left[\begin{array}{c} \mu \\ A_0 \mu \end{array}\right], \left[\begin{array}{cc} \Sigma & \Sigma A_0^T \\ A_0 \Sigma & A_0 \Sigma A_0^T \end{array}\right]\right)$$

- Bayesian updating of the normal distribution yields

$$\left(y \mid x_0 = x^p\right) \sim N\left(E\left[y \mid x_0 = x^p\right], Var\left[y \mid x_0\right]\right)$$

where

$$E\left[y \mid x_0 = x^p\right] = \mu + \Sigma A_0^T \left(A_0 \Sigma A_0^T\right)^{-1} \left(x^p - A_0 \mu\right)$$

or since a set of transactions, y^p, consistent with the financial statements, x^p, can be found such that $A_0 y^p = x^p$, we can replace x^p with $A_0 y^p$

$$\begin{aligned} E\left[y \mid x_0 = x^p\right] &= \mu + \Sigma A_0^T \left(A_0 \Sigma A_0^T\right)^{-1} \left(A_0 y^p - A_0 \mu\right) \\ &= \mu + \Sigma A_0^T \left(A_0 \Sigma A_0^T\right)^{-1} A_0 \left(y^p - \mu\right) \end{aligned}$$

and

$$Var\left[y \mid x_0\right] = \Sigma - \Sigma A_0^T \left(A_0 \Sigma A_0^T\right)^{-1} A_0 \Sigma$$

This is the result claimed above.

- In the special case $\Sigma = \sigma^2 I$, we have

$$\begin{aligned} E\left[y \mid x_0 = x^p\right] &= \mu + \sigma^2 I A_0^T \left(A_0 \sigma^2 I A_0^T\right)^{-1} \left(A_0 y^p - A_0 \mu\right) \\ &= \mu + A_0^T \left(A_0 A_0^T\right)^{-1} A_0 \left(y^p - \mu\right) \\ &= P_{R(A)} y^p + \left(I - P_{R(A)}\right) \mu \end{aligned}$$

and

$$\begin{aligned} Var\left[y \mid x_0\right] &= \sigma^2 I - \sigma^2 I A_0^T \left(A_0 \sigma^2 I A_0^T\right)^{-1} A_0 \sigma^2 I \\ &= \sigma^2 \left(I - A_0^T \left(A_0 A_0^T\right)^{-1} A_0\right) \\ &= \sigma^2 \left(I - P_{R(A)}\right) \end{aligned}$$

where $P_{R(A)} = A_0^T \left(A_0 A_0^T\right)^{-1} A_0$. Again, this is the result claimed above. It's time for an example.

Numerical example

Suppose we observe the following financial statements.

Balance sheets	Ending balance	Beginning balance
Cash	110	80
Receivables	80	70
Inventory	30	40
Property & equipment	110	100
Total assets	330	290
Payables	100	70
Owner's equity	230	220
Total equities	330	290

Income statement	for period
Sales	70
Cost of sales	30
SG&A	30
Net income	10

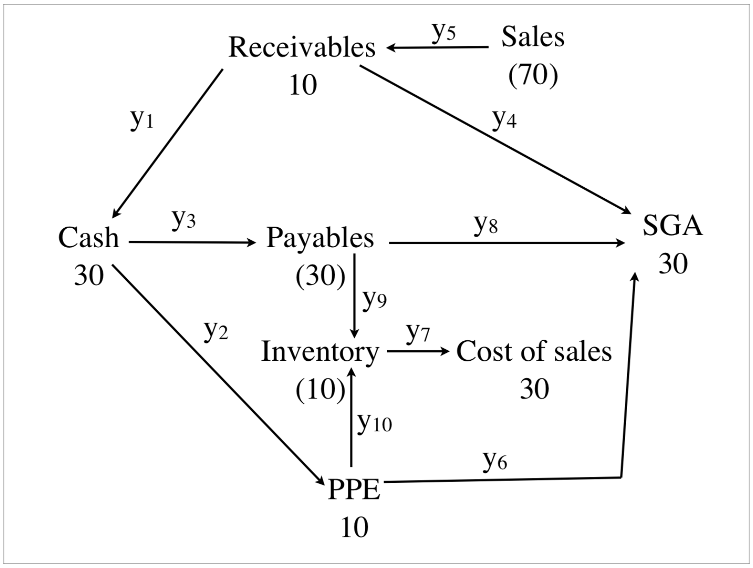
Let x be the change in account balance vector where credit changes are negative. The sum of x is zero; a basis for the left nullspace of A is a vector of ones.

change in account	amount
Δ cash	30
Δ receivables	10
Δ inventory	(10)
Δ property & equipment	10
Δ payables	(30)
sales	(70)
cost of sales	30
sg&a expenses	30

We envision the following transactions associated with the financial statements and are interested in recovering their magnitudes y.

transaction	amount
collection of receivables	y_1
investment in property & equipment	y_2
payment of payables	y_3
bad debts expense	y_4
sales	y_5
depreciation - period expense	y_6
cost of sales	y_7
accrued expenses	y_8
inventory purchases	y_9
depreciation - product cost	y_{10}

A crisp summary of these details is provided by a directed graph.



Directed graph of financial statements

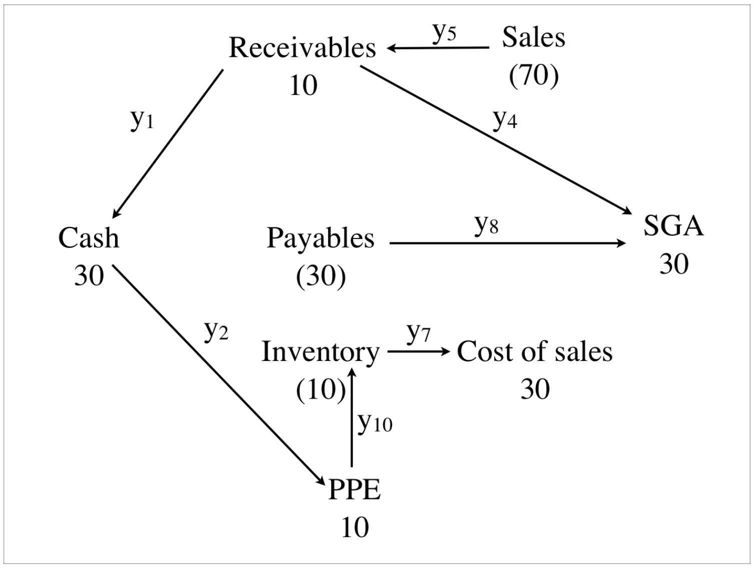
The A matrix associated with the financial statements and directed graph where credits are denoted by -1 is

$$A = \begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

and a basis for the nullspace is immediately identified by any set of linearly independent loops in the graph, for example,

$$N = \begin{bmatrix} 1 & 0 & 1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 1 & -1 \end{bmatrix}$$

A consistent solution y^p is readily identified by forming a spanning tree and solving the remaining transactions. For instance, let $y_3 = y_6 = y_9 = 0$, the spanning tree is depicted below



Spanning tree

.

Then, $(y^p)^T = \begin{bmatrix} 60 & 30 & 0 & 0 & 70 & 0 & 30 & 30 & 0 & 20 \end{bmatrix}$.

Now, suppose background knowledge $\Im$ regarding transactions is described by the first two moments

$$E\left[y^T \mid \Im\right] = \mu^T = \begin{bmatrix} 60 & 20 & 25 & 2 & 80 & 5 & 40 & 10 & 20 & 15 \end{bmatrix}$$

and

$$Var\left[y \mid \Im\right] = \Sigma = \begin{bmatrix} 10 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0.2 & 0 & 0 & 0 & 0.2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0.2 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.1 & 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0.1 & 10 & 0 & 3.5 & 0 & 0 & 0 \\ 0 & 0.2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3.5 & 0 & 5 & 0 & 0.2 & 0 \\ 0 & 0 & 0.2 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.2 & 0 & 1 & 0 \\ 0 & 0.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

maximum entropy priors for transactions are normally distributed with parameters described by the above moments.

Given financial statements x and background knowledge $\Im$, posterior beliefs regarding transactions are normally distributed with $E\left[y^T \mid x, \Im\right] =$

$$\begin{bmatrix} 58.183 & 15.985 & 12.198 & 1.817 & 70 & 5.748 & 30 & 22.435 & 19.764 & 0.236 \end{bmatrix}$$

and $Var\left[y \mid x, \Im\right] =$

$$\begin{bmatrix}
0.338 & 0.172 & 0.167 & -0.338 & 0 & 0.164 & 0 & 0.174 & -0.007 & 0.007 \\
0.172 & 0.482 & -0.310 & -0.172 & 0 & 0.300 & 0 & -0.128 & -0.182 & 0.182 \\
0.167 & -0.310 & 0.477 & -0.167 & 0 & -0.135 & 0 & 0.302 & 0.175 & -0.175 \\
-0.338 & -0.172 & -0.167 & 0.338 & 0 & -0.164 & 0 & -0.174 & 0.007 & -0.007 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.164 & 0.300 & -0.135 & -0.164 & 0 & 0.445 & 0 & -0.281 & 0.145 & -0.145 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.174 & -0.128 & 0.302 & -0.174 & 0 & -0.281 & 0 & 0.455 & -0.153 & 0.153 \\
-0.007 & -0.182 & 0.175 & 0.007 & 0 & 0.145 & 0 & -0.153 & 0.328 & -0.328 \\
0.007 & 0.182 & -0.175 & -0.007 & 0 & -0.145 & 0 & 0.153 & -0.328 & 0.328
\end{bmatrix}$$

As our intuition suggests, the posterior mean of transactions is consistent with the financial statements, $A\left(E\left[y \mid x, \Im\right]\right) = x$, and there is no residual uncertainty regarding transactions that are not in loops, sales and cost of sales are $y_5 = 70$ and $y_7 = 30$, respectively.