Actuarial Risk Matrices: The Nearest Positive Semidefinite Matrix Problem.

Dr. Adrian O'Hagan, Stefan Cutajar and Dr Helena Smigoc

School of Mathematics and Statistics
University College Dublin
Ireland

adrian.ohagan@ucd.ie

April, 2016

Features of a valid correlation matrix

- Correlation matrices:

  Diagonal elements all equal 1
  Matrix is symmetric
  All off-diagonal elements between 1 and -1 inclusive.

- A less intuitive property is that a correlation matrix must also be positive semidefinite:

$$\sum_i \sum_j a_i a_j \mathrm{Corr}(i,j) \geq 0 \quad \forall a_i, a_j \in \mathbf{R}.$$

  The variance of a weighted sum of random variables must be nonnegative for all choices of real weights.

How non-positive semidefinite matrices arise.

Reasons why a correlation matrix may not be positive semidefinite:

- Noise
- Elements estimated from disparate models
- Elements subjectively adjusted (to confer financial prudence, for example)
- Rounding
- Incomplete data/data with many outliers
- Correlation coefficients computed using inconsistent approaches (Pearson vs Spearman)

Norms

Starting matrix: A

Solution matrix: X

- Chebychev (maximum) norm:

$$\|A - X\|_{max} = max|A_{ij} - X_{ij}|$$

- Frobenius norm:

$$\|A - X\|_F = \sqrt{\sum_{i,j=1}^{n} (A_{ij} - X_{ij})^2}$$

Toy matrix.

Toy 5×5 correlation matrix with off-diagonal blocks of constants.

$$A = \begin{pmatrix} 1 & 0.5886 & -0.0292 & -0.0292 & -0.0292 \\ 0.5886 & 1 & -0.0292 & -0.0292 & -0.0292 \\ -0.0292 & -0.0292 & 1 & 0.8267 & -0.6952 \\ -0.0292 & -0.0292 & 0.8267 & 1 & -0.1146 \\ -0.0292 & -0.0292 & -0.6952 & -0.1146 & 1 \end{pmatrix}$$

Toy solution matrix

$$A = \begin{pmatrix} 1 & 0.5886 & -0.0292 & -0.0292 & -0.0292 \\ 0.5886 & 1 & -0.0292 & -0.0292 & -0.0292 \\ -0.0292 & -0.0292 & 1 & 0.8267 & -0.6952 \\ -0.0292 & -0.0292 & 0.8267 & 1 & -0.1146 \\ -0.0292 & -0.0292 & -0.6952 & -0.1146 & 1 \end{pmatrix}$$

Using the Alternating Projections Method (minimizing the Frobenius Norm) without off-diagonal constraints:

$$X = \begin{pmatrix} 1 & 0.5886 & -0.0289 & -0.0295 & -0.0290 \\ 0.5886 & 1 & -0.0289 & -0.0295 & -0.0290 \\ -0.0289 & -0.0289 & 1 & 0.8101 & -0.6819 \\ -0.0295 & -0.0295 & 0.8101 & 1 & -0.1244 \\ -0.0290 & -0.0290 & -0.6819 & -0.1244 & 1 \end{pmatrix}$$

Frobenius distance 0.0331 Chebychev distance 0.0133

Toy solution matrix

$$A = \begin{pmatrix} 1 & 0.5886 & -0.0292 & -0.0292 & -0.0292 \\ 0.5886 & 1 & -0.0292 & -0.0292 & -0.0292 \\ -0.0292 & -0.0292 & 1 & 0.8267 & -0.6952 \\ -0.0292 & -0.0292 & 0.8267 & 1 & -0.1146 \\ -0.0292 & -0.0292 & -0.6952 & -0.1146 & 1 \end{pmatrix}$$

Using the Alternating Projections Method (minimizing the Frobenius Norm) with off-diagonal block constraints:

$$X = \begin{pmatrix} 1 & 0.5886 & -0.0291 & -0.0291 & -0.0291 \\ 0.5886 & 1 & -0.0291 & -0.0291 & -0.0291 \\ -0.0291 & -0.0291 & 1 & 0.8101 & -0.6819 \\ -0.0291 & -0.0291 & 0.8101 & 1 & -0.1245 \\ -0.0291 & -0.0291 & -0.6819 & -0.1245 & 1 \end{pmatrix}$$

Frobenius distance 0.0331+ Chebychev distance 0.0133+

Current TMK approach: Igloo and ReMetrica

- Iterate between Igloo and ReMetrica.
- Igloo is generally very accurate in terms of the nearest PSD matrices identified.
- However Igloo unable to achieve the desired off-diagonal block structure.
- ReMetrica is able to incorporate the off diagonal block structure but is relatively inaccurate in producing “near” matrices.
- Iterative approach is slow and requires significant manual input.

Semidefinite programming

- Challenge can be written as an optimization problem with a linear objective function (minimizing a norm).
- Once the problem is identified to be a semidefinite programming problem there are several algorithms available.
- However they revolve around setting up constraints on all elements in the correlation matrix (PSD matrix, diagonal elements of 1, symmetry and off-diagonal blocks of constants).

Semidefinite programming

- Higham (2001) concludes that in order to compute the nearest correlation matrix for the classical problem (no off-diagonal blocks) we require $\frac{1}{2}n^4 + \frac{3}{2}n^2 + n + 1$ constraints.
- This is slow for very large n (but can be done, see for example MOSEK package in Matlab).
- The complication of having fixed off-diagonal blocks adds a considerable amount of additional constraints and hence would require an even greater increase in execution time.

Alternating projections method

- Positive semidefinite matrices (set $\mathcal{S}$): classical result tells us how to find a matrix that is positive semidefinite and closest to a given symmetric matrix A in the Frobenius norm:

$$A = M'DM$$

- M is an orthogonal matrix. D is a diagonal matrix.
- If A is not positive semidefinite some of the diagonal entries of D are negative.
- Let D_0 be a matrix obtained from D by setting all the negative entries in D equal to 0.
- Now $A_0 = M'D_0M$ is positive semidefinite and in Frobenius norm closest to A.

Alternating projections method

- Matrices with all the diagonal elements equal to 1 (set $\mathcal{U}$):

  if A_0 does not have all the diagonal entries equal to 1, set all the diagonal entries equal to 1.

- Matrices with diagonal elements equal to 1 AND with blocks of constants (set $\mathcal{V}$):

  if A_0 does not have all the diagonal entries equal to 1, set all the diagonal entries equal to 1.

  if A_0 does not have all its entries in a given block equal, compute the average of the entries of A in this block and put all entries in the block equal to the average.

Alternating projections method

- Assume that you want to find a matrix that is the closest to a given matrix A and is contained in the intersection of sets $\mathcal{S}$ and $\mathcal{V}$:
- $\mathcal{S}$: PSD matrices
  $\mathcal{V}$: matrices with diagonal elements equal to 1 and off-diagonal blocks of constants.
- We know (separately) how to find a closest point in $\mathcal{S}$ and how to find a closest point in $\mathcal{V}$
- But we don't know how to simultaneously find a closest point in the intersection of $\mathcal{S}$ and $\mathcal{V}$.
- Hence we ALTERNATE between the two PROJECTIONS...

Alternating projections method

- Hence we ALTERNATE between the two PROJECTIONS

  $P_{\mathcal{S}}(A)$

  $P_{\mathcal{V}}(P_{\mathcal{S}}(A))$

  $P_{\mathcal{S}}(P_{\mathcal{V}}(P_{\mathcal{S}}(A)))$ ...

- If this process converges, the dual objectives are satisfied (typical convergence criterion is that maximum individual element change between two successive iterations is less than 5×10^{-5}).
- Make sure to terminate the algorithm on a matrix projection into $\mathcal{S}$!!
- Some harder math: Dykstra's projection algorithm to guarantee convergence.

Results: Alternating projections method

APM1: classical approach (ignores off-diagonal blocks of constants).
APM2: preserves off-diagonal blocks of constants.

Table : Comparison of Results on Sample Matrix A_1: dimension 155×155

	$\min eig(X_1)$	$\|A_1 - X_1\|_F$	$\|A_1 - X_1\|_{max}$	Time
TMK	$-3.05E-16$	1.0528	0.038	$\approx$ 4 hours
APM1	$1.00E-07$	0.6756	0.0415	0.2064 s
APM2	$1.00E-07$	0.7956	0.0468	3.204 s

Minimizing Chebychev norm

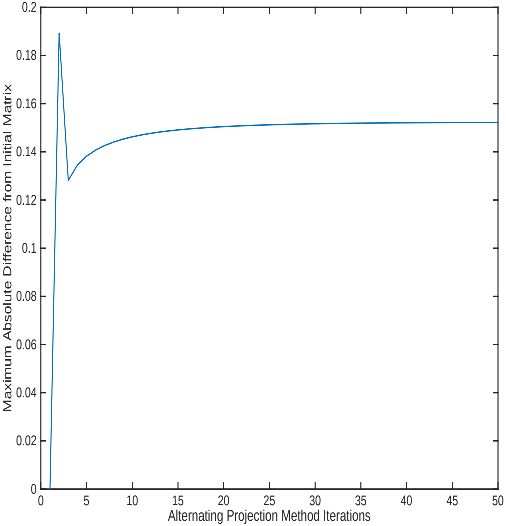
- APM described above gives an optimal solution in the Frobenius norm fairly quickly, so that problem is satisfactorily addressed.
- However there still remains the outstanding question: how to deal with other matrix norms, most notably the Chebychev norm ?
- Least Maximum Norm algorithm is one possibility (Fmincom package in MATLAB) but as with semidefinite programming is very slow.
- Instead we try to optimise for Chebychev norm using the (much faster) tools already developed (and one new one)

Minimizing Chebychev norm within APM: approach 1 - the crude method

- Approach 1 (crude): record the maximum (Chebychev) norm at each iteration of the APM.

- The minimal Chebychev norm among all the matrices produced in the APM iterations will typically occur before convergence to the minimal Frobenius norm.

Minimizing Chebychev norm within APM: approach 1 - the crude method



Minimizing Chebychev norm within APM: approach 2 - shrinking method

- A convex combination of our original matrix A (perfect in Chebychev norm, but not positive definite) and some positive definite matrix B:

$$C(t) = (1 - t)A + tB$$

  $t = 0$: just returns the original matrix $C(0) = A$.
  $t = 1$: guaranteed positive definite matrix $C(1) = B$.
- There exists (a minimal) t^* in $(0, 1)$ such that $C(t)$ is positive definite for all $t > t^*$.
- $C(t^*)$ is the closest (in any norm) positive semidefinite matrix to A among all matrices of the form $(1 - t)A + tB$.
- Two challenges: how to find t^* and what to use for B.

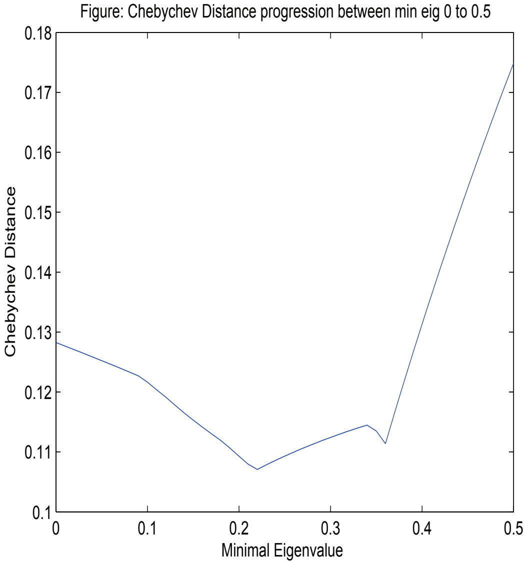
The shrinking method: finding t^* - the bisection method

- $C(0)$ is not PSD and $C(1)$ is.
- Check if $C(1/2)$ is PSD.
- If $C(1/2)$ is PSD check if $C(1/4)$ is PSD.
  If $C(1/4)$ is PSD check $C(1/8)$, otherwise check $C(3/8)$...
- If $C(1/2)$ is not PSD check if $C(3/4)$ is PSD.
  If $C(3/4)$ is PSD check $C(5/8)$, otherwise check $C(7/8)$...

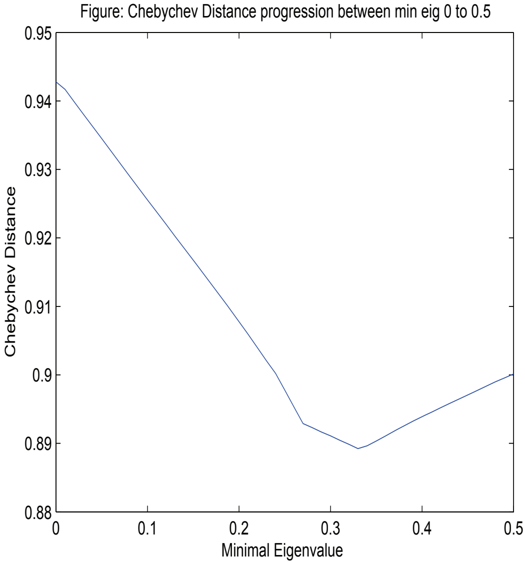
The shrinking method: finding B

- We want B to be close to A and PSD, can do this using APM.
- But now we no longer need (or want) the minimal eigenvalue to be 0.
- Exploit this by gradually increasing the minimal eigenvalue of B and recording the maximum (Chebychev) norm for the answer.
- In this way an optimal minimal eigenvalue and accompanying B is identified.

Results: shrinking method



Results: shrinking method



Minimizing Chebychev norm within APM: approach 3 ??

- Join Approach 1 (checking Chebychev norm values as APM progresses) and Approach 2 (Shrinking method):
- Set B to be the matrix that is the closest in the Chebyshev norm to A in the iterative process for some given eigenvalue λ
- Then as before increase λ.

$$C(t) = (1 - t)A + tB$$

Another open idea...

- Not all risks are created equal.
- Large positive correlations in the starting matrix A point to risk pairings that are more inclined to simultaneously materialise in large losses.
- Amend the algorithms to prioritise “nearness” among these cell pairings.
- Could have a weighted Frobenius norm with higher weights on positive values in A...
- ...or an adapted Chebychev norm that only looks at maximum deviations from positive values in A.

Conclusions

- All methods work well
- The Alternating Projections Method (APM) is readily applicable and is optimal in terms of convergence speed.
- APM has linear convergence rate, but still very efficient.
- APM minimizes Frobenius norm - must track Chebychev norm and corresponding matrices if minimization of latter is the goal.
- The Semidefinite Progrmaming (SDP) method proved to be as accurate as the APM (same resultant matrices).
- However SDP significantly slower than APM.
- Shrinking method appears to hold promise for minimizing Chebychev norm when using APM.

Upcoming talks and publication

- Our paper will be under review and a draft available on ArXiV soon !
- Planned presentations at:

  Society of Actuaries in Ireland
  Tomorrow !

  51st Actuarial Research Conference (Society of Actuaries)
  University of Minneapolis, July 2016.

  GIRO 2016 (Institute and Faculty of Actuaries)
  Dublin, September 2016.

- Updated slides available on request.

Thank you

Mr Brian Heffernan and Mr Tetsushi Imatomi, TokioMarine Kiln.

Questions and Discussion