Manchester Numerical Analysis

Anderson Acceleration of the Alternating Projections Method for Computing the Nearest Correlation Matrix

Nick.Higham@manchester.ac.uk   Natasa.Strabic@manchester.ac.uk

Fixed-point Iteration

Seek x_* such that $g(x_*) = x_*$ for $g: \mathbb{R}^n \to \mathbb{R}^n$:

$$x_{k+1} = g(x_k), \quad k \geq 1, \qquad x_0 \in \mathbb{R}^n \text{ given.}$$

Anderson Acceleration (basic)

The iteration history for some chosen $m \geq 0$ is

$$\begin{matrix} \dots & x_{k-m} & x_{k-(m-1)} & \dots & x_{k-1} & x_k \\ \dots & g(x_{k-m}) & g(x_{k-(m-1)}) & \dots & g(x_{k-1}) & g(x_k). \end{matrix}$$

- Define x_{k+1} using all this information; [Anderson 1965].

Given $x_0 \in \mathbb{R}^n$ and $m \geq 1$.

$x_1 = g(x_0)$.
For $k = 1, 2 \dots$
  $m_k = \min\{m, k\}$.
  Compute $\theta^{(k)}$ solving $\min_{\theta \in \mathbb{R}^{m_k}} \|v_k - u_k\|_2^2$, where

$$u_k = x_k + \sum_{j=1}^{m_k} \theta_j (x_{k-j} - x_k),$$

$$v_k = g(x_k) + \sum_{j=1}^{m_k} \theta_j \left(g(x_{k-j}) - g(x_k)\right).$$

  $x_{k+1} = v_k$ with the parameters from $\theta^{(k)}$.
end

- If g linear and $m_k = k$, essentially GMRES.

Anderson Acceleration (practical)

- Equivalent version designed for $f(x_*) = 0$; [Fang, Saad 2009], [Walker, Ni 2011].

$\mathcal{X}_k = [\Delta x_{k-m_k} \ \cdots \ \Delta x_{k-1}] \in \mathbb{R}^{n \times m_k}$, $\Delta x_i = x_{i+1} - x_i$,
$\mathcal{G}_k = [\Delta g_{k-m_k} \ \cdots \ \Delta g_{k-1}] \in \mathbb{R}^{n \times m_k}$, $\Delta g_i = g(x_{i+1}) - g(x_i)$,
$\mathcal{F}_k = \mathcal{G}_k - \mathcal{X}_k$, $f_k = g(x_k) - x_k$.

Given $x_0 \in \mathbb{R}^n$ and $m \geq 1$.
$x_1 = g(x_0)$.
For $k = 1, 2 \dots$
  $m_k = \min\{m, k\}$.
  Compute $\gamma^{(k)}$ which solves $\min_{\gamma \in \mathbb{R}^{m_k}} \|f_k - \mathcal{F}_k \gamma\|_2^2$.
  $x_{k+1} = g(x_k) - \mathcal{G}_k \gamma^{(k)}$.
end

LS Problem Details

At each iteration first col $\mathcal{F}_k$ is removed and new last column added $\Rightarrow$ update QR factors of $\mathcal{F}_k$ as iteration progresses.

Nearest Correlation Matrix (NCM)

A **correlation matrix** is pos semidef with unit diagonal.

Given $A = A^T \in \mathbb{R}^{n \times n}$, compute

$$\min\{\, \|A - X\|_F : X \text{ a c.m.} \,\}, \tag{1}$$

$$\min\{\, \|A - X\|_F : X \text{ a c.m.}, \lambda_{\min}(X) \geq \delta,\ 0 < \delta \leq 1 \,\}, \tag{2}$$

$$\min\{\, \|A - X\|_F : X \text{ a c.m.}, x_{ij} = a_{ij} \text{ for } (i,j) \in \mathcal{N} \,\}. \tag{3}$$

For (1) and (2) solution is unique, for (3) unique if it exists.

Applications Include

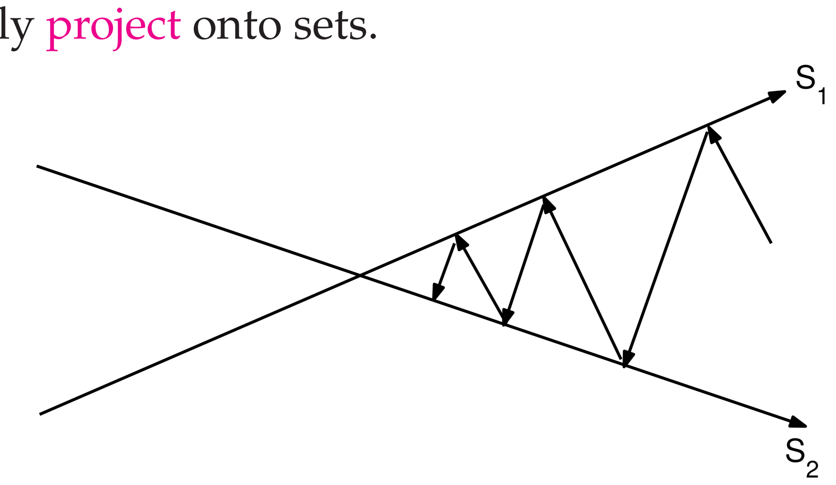
- Portfolio risk management in finance.
- Prediction of electricity peak-demand.
- Modelling public health.
- Simulation of wireless links in vehicular networks.
- Analysis of carbon dioxide storage resources in the US.
- Determination of insurance premiums for crops.
- Genetic evaluations for thoroughbred horse breeding.

Alternating Projections Method

- First method with guaranteed convergence (at best linear); [Higham 2002].
- Widely implemented (MATLAB, Python, R, SAS).
- Solves all three problems.

NCM is the nearest matrix to A in the intersection of S_1 = symm pos semidef and S_2 = symm unit diag.

Iteratively project onto sets.



Convergence results: [von Neumann 1933] – for subspaces. [Dykstra 1983], [Boyle, Dykstra 1986] – corrections for closed convex sets.

Newton Method

- Global quadratic convergence; [Qi, Sun 2006].
- In NAG library with improvements [Borsdorf, Higham 2010].
- Cannot solve the fixed elements variant.

Motivation for NCM problems

Correlation matrix obtained in practice might be

- **Indefinite**, as a result of
  - Missing data.
  - Expert judgement.
  - Asynchronous data.
  - Aggregation methods.
- **Singular**: always when there are fewer observations than variables.

Alternating Projections for NCM

A fixed-point variant of the code:

$\Delta_0 = 0$, $Y_0 = A$
For $k = 1, 2, \dots$
  $(Y_k, \Delta_k) = g(Y_{k-1}, \Delta_{k-1})$
end
Return Y_k.

where the function g is $(Y_k, \Delta_k) = g(Y_{k-1}, \Delta_{k-1})$

$R_k = Y_{k-1} - \Delta_{k-1}$
$X_k = \mathcal{P}_{S_1}(R_k)$
$\Delta_k = X_k - R_k$
$Y_k = \mathcal{P}_{S_2}(X_k)$.

Individual projections:

- $\mathcal{P}_{S_1}(A)$: set all $\lambda_i(A) < \delta$ to δ, where $0 \leq \delta \leq 1$.
- $\mathcal{P}_{S_2}(A)$: overwrite with corresponding fixed elements.

Stopping test: $\|Y_k - X_k\|_F \leq \text{tol}\|Y_k\|_F$.

Anderson Acceleration for NCM

To use NCM code with Anderson acceleration write matrices as vectors using the vec operator.

Timings

Computing positive definite NCM with $\delta = 0.1$ (for Anderson acceleration, $m = 2$) for six RiskMetrics matrices of order 387.

Matrix	nearcorr		nearcorr_AA			
	it	t	itAA	t	t_apm	t_AA
1	1410	20.50	383	10.77	5.70	3.12
2	2100	33.93	513	15.83	8.52	4.56
3	1900	31.14	414	11.58	5.97	3.54
4	1586	29.06	369	12.83	7.09	3.54
5	1812	31.30	400	12.99	7.16	3.62
6	1794	29.08	393	11.63	6.20	3.40

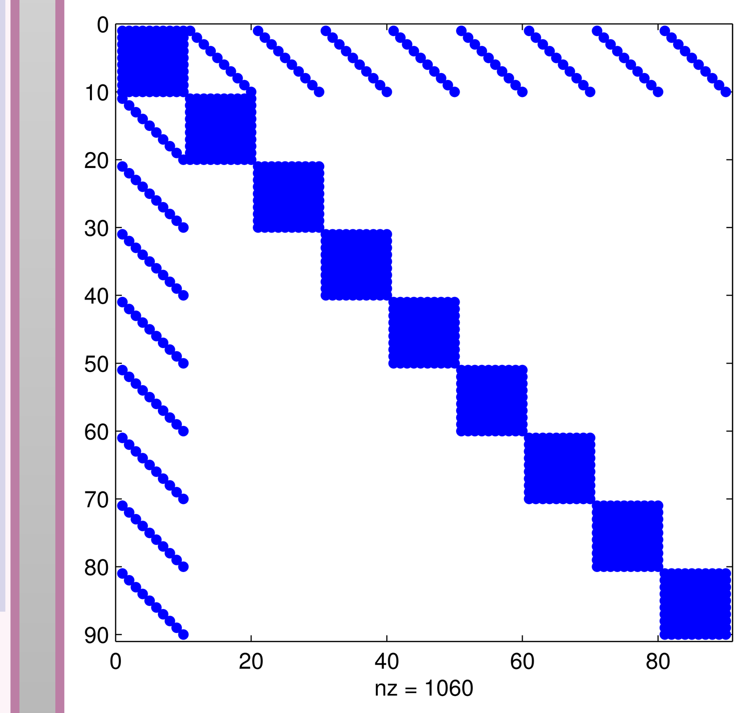
- Speedups 2–3.
- LS solution is a minor cost.

Small Examples

n	it	itAA					
		$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$
4	39	15	10	9	9	9	9
5	27	17	14	12	11	10	10
6	801	305	212	117	126	40	31
7	33	15	10	10	10	9	9

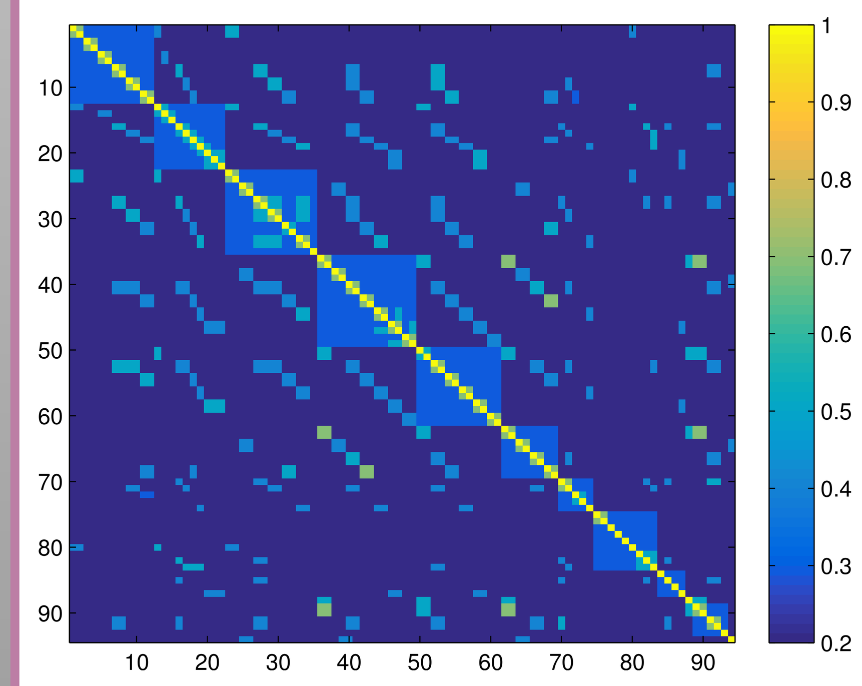
- Anderson acceleration reduces number of iterations for alternating projections method by **factor ≈ 3–25**.
- Small values of m are sufficient.

Fixed Elements Examples



Keep fixed:

- (1,1) block.
- Main diagonal (not unit).
- "Small" diagonals.



Diag. blocks are fixed.

n	it	it_fe	itAA_fe				
			$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
90	29	169	93	70	55	45	39
94	18	40	15	14	12	12	12

- Additional constraints increase the iteration count.
- **Anderson acceleration remains effective.**
- For the variants, iterations reduce by **at least a third**.

Reference

N. J. Higham and N. Strabić. *Anderson acceleration of the alternating projections method for computing the nearest correlation matrix.* MIMS EPrint 2015.39, August 2015. 22 pp.