The convergence of stationary iterations with indefinite splitting

Michael C. Ferris

Joint work with: Tom Rutherford and Andy Wathen

University of Wisconsin, Madison

6th International Conference on Complementarity Problems
Berlin, Germany
August 7, 2014

The problems

- VI(F, C): $x^* \in C$, $\langle F(x^*), x - x^* \rangle \geq 0, \forall x \in C$
- MOPEC: x^*, y:

$$x_i^* \text{ solves } \min_{x_i \in K_i(x^*_{-i}, y)} \theta(x_i, x^*_{-i}, y), \forall i$$

$$y \text{ solves } VI(F(x^*, \cdot), C)$$

$$\begin{bmatrix} \mathcal{A}_1 & A_{1,2} & \cdots & A_{1,p} & E_1 \\ A_{2,1} & \mathcal{A}_2 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & A_{p-1,p} & E_{p-1} \\ A_{p,1} & \cdots & A_{p,p-1} & \mathcal{A}_p & E_p \\ F_1 & \cdots & F_{p-1} & F_p & D \end{bmatrix}$$

Strongly Convex Nash Equilibria

$$\min_{x_1 \geq 0} \; \frac{1}{2}x_1^2 - \theta x_1 x_2 - 4x_1 \text{ s.t. } x_1 + x_2 \geq 1$$
$$\min_{x_2 \geq 0} \; \frac{1}{2}x_2^2 - x_1 x_2 - 3x_2$$

- No solution for $\theta \geq 1$:

$$x_1(x_2) = (\theta x_2 + 4)_+, \quad x_2(x_1) = (x_1 + 3)_+$$

- Solution $-\frac{4}{3} \leq \theta < 1$: $x_1 = \frac{4+3\theta}{1-\theta}$, $x_2 = x_1 + 3$
- Solution $\theta \leq -\frac{4}{3}$: $x_1 = 0$, $x_1 = 3$
- Jacobi works provided $\theta < 1$, but theory fails

The Issues

This is not the optimality conditions of a single optimization problem:

$$0 \leq \left[\begin{array}{cc|c} 1 & 1 & -\theta \\ 1 & 0 & 1 \\ \hline -1 & & 1 \end{array}\right] \begin{bmatrix} x_1 \\ -p_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} \quad \perp \quad \begin{bmatrix} x_1 \\ -p_1 \\ x_2 \end{bmatrix} \geq 0$$

- The matrix $\mathcal{A}$ in general is never diagonally dominant except in trivial cases
- Iterations based on succesive inversion of local blocks (or successive optimization of local strategies) can converge.
- We establish sufficient conditions which guarantee convergence of block Jacobi and block Gauss-Seidel iterations for such matrices.

Iteration with Indefinite Splitting

$$Ax = b$$

Splitting $A = P - N$ naturally leads to a stationary iteration of the form

$$x_0 \text{ arbitrary}, \quad Px_{k+1} = Nx_k + b, \quad k = 0, 1, \ldots$$

- This iteration may or may not converge; simply applicable sufficient conditions for convergence are particularly valuable.
- Most well-known such conditions are diagonal dominance:
  - if the preconditioner is $P = \mathrm{diag(A)}$ (leading to Jacobi iteration) or
  - P is the lower triangular part of A (leading to Gauss-Seidel iteration),

  then convergence is guaranteed if the strict diagonal dominance condition

$$|a_{i,i}| > \sum_{j=1,\ldots,n, j\neq i} |a_{i,j}|, \quad i = 1, \ldots, n \tag{1}$$

is satisfied by $A = \{a_{i,j}, i, j = 1, \ldots, n\}$.

Weaker diagonal dominance conditions

For irreducible matrices, it is well documented that the weaker condition

$$|a_{i,i}| \geq \sum_{j=1,\ldots,n, j\neq i} |a_{i,j}|, \quad i = 1, \ldots, n \tag{2}$$

is also sufficient provided strict inequality holds for at least one row index, i
The condition (1) or (2) also guarantees that $A \in \mathbb{R}^{n\times n}$ is invertible, so a unique solution exists.

The Setting

We focus on matrices of the form

$$\mathcal{A} = \begin{bmatrix} \mathcal{A}_1 & A_{1,2} & \cdots & A_{1,p} & E_1 \\ A_{2,1} & \mathcal{A}_2 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & A_{p-1,p} & E_{p-1} \\ A_{p,1} & \cdots & A_{p,p-1} & \mathcal{A}_p & E_p \\ F_1 & \cdots & F_{p-1} & F_p & D \end{bmatrix} \quad (3)$$

where

$$\mathcal{A}_i = \begin{bmatrix} Q_i & B_i^T \\ B_i & 0 \end{bmatrix}, i = 1, \ldots, p \quad (4)$$

with $Q_i = Q_i^T \in \mathbb{R}^{n_i \times n_i}$ positive definite and $B_i \in \mathbb{R}^{m_i \times n_i}$ of full rank $m_i < n_i$ for each i ($m_i > 0$). These conditions guarantee that each $\mathcal{A}_i$ is invertible. The submatrix $D \in \mathbb{R}^{s \times s}, s \geq 0$ must be symmetric and invertible (unless $s = 0$).

Existing Block Theory

For the blocked matrix (3) a result of Feingold and Varga (1962) applies: If $\mathcal{A}$ is block irreducible and

$$
\begin{aligned}
(\|\mathcal{A}_i^{-1}\|_2)^{-1} &\geq \|E_i\|_2 + \sum_{j=1,\ldots,p,j\neq i} \|A_{i,j}\|_2, \quad i = 1,\ldots,p \quad (5) \\
\text{and} \quad (\|D^{-1}\|_2)^{-1} &\geq \sum_{j=1,\ldots,p,j\neq i} \|F_i\|_2 \quad (6)
\end{aligned}
$$

with strict inequality in (6) or for at least one index, i, in (5), then $\mathcal{A}$ is invertible (existence and uniqueness)

Relation to Iteration

Before considering these conditions in more detail, consider a block Jacobi or block Gauss-Seidel iteration based on the splitting with

$$P = \begin{bmatrix} \mathcal{A}_1 & & & & \\ & \mathcal{A}_2 & & & \\ & & \ddots & & \\ & & & \mathcal{A}_p & \\ & & & & D \end{bmatrix} \quad \text{or} \quad P = \begin{bmatrix} \mathcal{A}_1 & & & & \\ A_{2,1} & \mathcal{A}_2 & & & \\ \vdots & \ddots & \ddots & & \\ A_{p,1} & \cdots & A_{p,p-1} & \mathcal{A}_p & \\ F_1 & \cdots & F_{p-1} & F_p & D \end{bmatrix}$$

Asymptotic convergence of the corresponding stationary (or simple) iteration will be guaranteed for any starting vector if all of the eigenvalues, λ, of $I - P^{-1}\mathcal{A}$ lie strictly inside the unit disc.

The link

Such eigenvalues satisfy $(I - P^{-1}\mathcal{A})x = \lambda x, x \neq 0$ or equivalently $(\mathcal{A} + (\lambda - 1)P)x = 0, x \neq 0$. In the case of block Jacobi, asymptotic convergence will be guaranteed if there does not exist any λ with $|\lambda| \geq 1$ such that the matrix

$$\mathcal{A}(\lambda) = \mathcal{A} + (\lambda - 1)P = \begin{bmatrix} \lambda\mathcal{A}_1 & A_{1,2} & \cdots & A_{1,p} & E_1 \\ A_{2,1} & \lambda\mathcal{A}_2 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & A_{p-1,p} & E_{p-1} \\ A_{p,1} & \cdots & A_{p,p-1} & \lambda\mathcal{A}_p & E_p \\ F_1 & \cdots & F_{p-1} & F_p & \lambda D \end{bmatrix}$$

is singular. But

$$(\|(\lambda\mathcal{A}_i)^{-1}\|_2)^{-1} = |\lambda|(\|\mathcal{A}_i^{-1}\|_2)^{-1} \geq (\|\mathcal{A}_i^{-1}\|_2)^{-1}$$

whenever $|\lambda| \geq 1$ with a identical argument holding for D. Hence satisfaction of the conditions (5),(6) not only guarantees invertibility of $\mathcal{A}$, but also guarantees covergence of the block Jacobi iteration.

Another piece

Let μ_i denote the smallest eigenvalue of the positive definite matrix Q_i and γ_i denote the smallest eigenvalue of the positive definite (Schur complement) matrix $B_i Q_i^{-1} B_i^T$, then there are no eigenvalues of

$$\mathcal{A}_i = \begin{bmatrix} Q_i & B_i^T \\ B_i & 0 \end{bmatrix}$$

in the interval

$$\left(\frac{1}{2} \left(\mu_i - \sqrt{\mu_i^2 + 4\gamma_i \mu_i} \right), \mu_i \right)$$

which contains the origin.

Finally...

If the matrix $\mathcal{A}$ given by (3),(4) is block irreducible, then it is invertible and the block Jacobi and block Gauss-Seidel iterations for a linear system $\mathcal{A}x = b$ converge to x for any starting vector if

$$\min\left\{\frac{1}{2}\left(\sqrt{\mu_i^2 + 4\gamma_i\mu_i} - \mu_i\right), \mu_i\right\} \geq \|E_i\|_2 + \sum_{j=1,\ldots,p,j\neq i} \|A_{i,j}\|_2, \quad i = 1,\ldots,p \tag{7}$$

$$\text{and} \quad d \geq \sum_{j=1,\ldots,p,j\neq i} \|F_i\|_2 \tag{8}$$

with strict inequality in (8)[1] or for at least one index, i, in (7).

[1] d is the absolute value of eighenvalue of D closest to origin

A Simplification

If for each $i = 1, \ldots, p$, $\gamma_i \geq 2\mu_i$ then $\mathcal{A}$ is invertible and the block Jacobi and block Gauss-Seidel iterations for a linear system $\mathcal{A}x = b$ converge to x for any starting vector if

$$\mu_i \geq \|E_i\|_2 + \sum_{j=1,\ldots,p,j\neq i} \|A_{i,j}\|_2, \quad i = 1, \ldots, p \tag{9}$$

$$\text{and} \quad d \geq \sum_{j=1,\ldots,p,j\neq i} \|F_i\|_2 \tag{10}$$

with strict inequality in (10) or for at least one index, i, in (9).

Simple example

$$\min_{x_1} \; 0.5x_1^2 - \theta x_1 x_2 - 4x_1 \text{ s.t. } 2x_1 + 0.5x_2 = 1$$

$$\min_{x_2} \; 0.5x_2^2 - x_1 x_2 - 3x_2$$

$$\left[\begin{array}{cc|c} 1 & 2 & -\theta \\ 2 & & 0.5 \\ \hline -1 & & 1 \end{array}\right] \begin{bmatrix} x_1 \\ -p_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

- Solution: $x_1 = -0.2, x_2 = 2.8,\; p_1 = -1.4\theta - 2.1$
- Jacobi works, but convergence guaranteed if $|\theta| < \sqrt{3}/2$.

Extensions

- Can also prove same result for SOR schemes
- Can apply regularization (proximal iterations) on the constraints: for $\epsilon_i, \alpha_i > 0$

$$\mathcal{A}_i = \begin{bmatrix} Q_i + \alpha_i I & B_i^T \\ B_i & -\epsilon_i I \end{bmatrix},$$

  can be used for some subset (or indeed all) of the indices $i = 1, \ldots, p$.
- This increases the value of μ_i and γ_i in the above and strengthens the theory
- No rates given here

Strongly convex optimization

$$\min_{x_1} \ \frac{1}{2}x_1^2 - x_1x_2 - 4x_1 \text{ s.t. } x_1 + x_2 = 1$$
$$\min_{x_2} \ \frac{1}{2}x_2^2 - x_1x_2 - 3x_2$$

$$\left[\begin{array}{cc|c} 1+\alpha_1 & 1 & -1 \\ 1 & -\epsilon_1 & 1 \\ \hline -1 & & 1 \end{array}\right] \begin{bmatrix} x_1 \\ -p_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

- Solution: $x_1 = -1, x_2 = 2,\ p_1 = -7$
- Jacobi fails: after 4 steps back at $(1,1)^T$
- Modified Jacobi $\alpha_1, \epsilon_1 = 0.1$ solves in 50 steps

Extension to Inequality Case (QP iterates)

- Each Jacobi iterate replaces system of equations with solution of a small(er) scale Quadratic Programs:

$$\min_{x_i \geq 0} \ \frac{1}{2} x_i^T Q_i x_i + c_i(x_{-i})^T x_i \text{ s.t. } B_i x_i = b(x_{-i})$$

- Solution is typically found in (many) fewer iterations than unconstrained case
- Can use any QP solver for subproblems (and/or VI solver)

Extension to Inequality Case (Normal Map)

- Replace systems of equations by (normal map formulation of) complementarity problems $\forall i$:

$$
\begin{aligned}
&Q_i((x_i)_+) - B_i^T p + c_i + x_i - (x_i)_+ = 0 \\
&B_i((x_i)_+) = b_i
\end{aligned}
$$

- Note this is a natural extension of the case considered above
- Choose active set at each iteration based on prediction from previous iteration
- Need to employ a regularization on the subproblem constraints
- Apply theory to all selections of the resulting linear systems

Extension to Inequality Case (Interior Point)

- Apply interior point code to solve each QP subproblem

$$\min_{x_i \geq 0} \; \frac{1}{2} x_i{}^T Q_i x_i + c_i(x_{-i})^T x_i \text{ s.t. } B_i x_i = b(x_{-i})$$

- Resulting systems to solve have form

$$\begin{bmatrix} Q_i + \bar{X}_i^{-1} \bar{W}_i & B_i^T \\ B_i & 0 \end{bmatrix} i = 1, \ldots, p$$

  where $\bar{X}$ and $\bar{W}$ are diagonals of iterates and slacks at previous iteration

- Update barrier parameter after each Jacobi step

Economic Application

- Model is a partial equilibrium, geographic exchange model.
- Goods are distinguished by region of origin.
- There is one unit of region r goods.
- These goods may be consumed in region r or they may be exported.
- Each region solves:

$$\min_{X,T_r} f_r(X,T) \text{ s.t. } F(X,T) = 0,\ T_j = \bar{T}_j, j \neq r$$

  where $f_r(X,T)$ is a quadratic form and $F(X,T)$ is linear and defines X uniquely as a function of T.
- $F(X,T)$ defines an equilibrium; here it is simply a set of equations, not a complementarity problem

Results

Gauss-Seidel residuals

Iteration	deviation
1	3.14930
2	0.90970
3	0.14224
4	0.02285
5	0.00373
6	0.00061
7	0.00010
8	0.00002
9	0.00000

Tariff revenue

region	SysOpt	MOPEC
1	0.117	0.012
2	0.517	0.407
3	0.496	0.214
4	0.517	0.407
5	0.117	0.012

- Note that competitive solution produces much less revenue than system optimal solution
- Model has non-convex objective, but each subproblem is solved globally (lindoglobal)

Conclusions

- MOPEC problems capture complex interactions between optimizing agents
- Policy implications addressable using MOPEC
- MOPEC available to use within the GAMS modeling system
- New sufficient conditions for existence, uniqueness and convergence shown in special cases
- Many new settings available for deployment; need for more theoretic and algorithmic enhancements