The Computation of the Group Inverse and related properties of Markov Chains via Perturbations

Jeffrey J Hunter

Professor Emeritus of Statistics, Massey University,

Professor of Mathematical Sciences, AUT University,

New Zealand

23rd International Workshop on Matrices and Statistics

Ljubljana, Slovenia, June 8 - 12, 2014

Outline

1. Introduction
2. Stationary distributions of Markov chains
3. Mean first passage times
4. Generalized matrix inverses
5. Solving systems of linear equations
6. The group inverse
7. G-inverses of Markovian kernels
8. Solving for stationary distributions
9. Solving for mean first passage times
10. Solving for the group inverse
11. Perturbation procedures
12. The algorithms
13. Computation comparisons

1. Introduction

Let P = $[p_{ij}]$ be the transition matrix of an irreducible, discrete time Markov chain (MC) $\{X_n\}$ ($n \geq 0$) with finite state space $S = \{1, 2, \ldots, m\}$.

$$i.e.\ \ p_{ij} = P\{X_n = j \mid X_{n-1} = i\} \text{ for all } i, j \in S.$$

We are interested in developing efficient ways of finding three key properties of such chains using perturbations:

(i) the stationary probabilites $\{\pi_j\}$, $(1 \leq j \leq m)$.

(ii) the mean first passage times $\{m_{ij}\}$,$(1 \leq i, j \leq m)$.

(iii) the group inverse of $I - P$, $A^{\#}$.

2. Stationary distributions

Let $\boldsymbol{\pi}^T = (\pi_1, \pi_2, \ldots, \pi_m)$ be the stationary prob. vector of the Markov chain.

We need to solve $\pi_j = \sum_{i=1}^{m} \pi_i p_{ij}$ with $\sum_{i=1}^{m} \pi_i = 1$,

i.e. $$\boldsymbol{\pi}^T (I - P) = \boldsymbol{0}^T \text{ with } \boldsymbol{\pi}^T \boldsymbol{e} = 1.$$

3. Mean first passage times

Let T_{ij} be the first passage time RV from state i to state j,

i.e. $T_{ij} = \min\{n \geq 1$ such that $X_n = j$ given that $X_0 = i\}$.

T_{ii} is the first return to state i.

Let $m_{ij} = E[T_{ij} \mid X_0 = i]$, the mean first passage time from state i to state j.

The mean first passage times

Let $M = [m_{ij}]$ be the matrix of mean first passage times

It is well known that

$$m_{ij} = 1 + \sum_{k \neq j} p_{ik} m_{kj},$$

with $m_{jj} = 1/\pi_j$.

M satisfies the matrix equation

$$(I - P)M = E - PD,$$

where $E = [1] = \mathbf{e}\mathbf{e}^T$, and

$D = M_d = [\delta_{ij} m_{ij}] = (\Pi_d)^{-1}$ (with $\Pi = \mathbf{e}\boldsymbol{\pi}^T$).

4. Generalized matrix inverses

A generalized inverse of a matrix A is any matrix A^{-} such that

$$AA^{-}A = A\,.$$

A^{-} is a “one condition” g-inverse

A^{-} is an “equation solving” g-inverse.

5. Solving systems of linear equations

A necessary and sufficient condition for

$$AXB = C$$

to have a solution is

$$AA^{-} CB^{-} B = C.$$

If this consistency condition is satisfied, the general solution is given by

$$X = A^{-}CB^{-} + W - A^{-}AWBB^{-}$$

where W is an arbitrary matrix.

(Penrose, 1955), (Rao,1955)

6. The group inverse

Let A be a square matrix with real elements, such that rank(A) = rank(A^2).

Then the matrix $A^\#$ which satisfies

Condition 1: $AA^\# A = A$
Condition 2: $A^\# AA^\# = A^\#$
Condition 5: $AA^\# = A^\# A$

exists, is unique, and is called the “group inverse” of A .

i.e. $A^\#$ is a 1-condition g-inverse with 2 additional conditions.

7. G-inverses of Markovian kernels

Let P be the transition matrix of a finite irreducible Markov chain with stationary probability vector $\boldsymbol{\pi}^T$.

Let $\mathbf{e}^T = (1, 1, \ldots, 1)$ and $\boldsymbol{t}$ and $\boldsymbol{u}$ be any vectors.

$I - P + \boldsymbol{tu}^T$ is non-singular $\Leftrightarrow$ $\boldsymbol{\pi}^T\boldsymbol{t} \neq 0$ and $\boldsymbol{u}^T\mathbf{e} \neq 0$.

$\boldsymbol{\pi}^T\boldsymbol{t} \neq 0$ and $\boldsymbol{u}^T\mathbf{e} \neq 0 \;\Rightarrow$

$[I - P + \boldsymbol{tu}^T]^{-1}$ is a g-inverse of $I - P$.

If G is any g-inverse of $I - P$, then $\exists$ vectors $\boldsymbol{f}$, $\boldsymbol{g}$, $\boldsymbol{t}$ and $\boldsymbol{u}$ with $\boldsymbol{\pi}^T\boldsymbol{t} \neq 0$ and $\boldsymbol{u}^T\mathbf{e} \neq 0$ such that

$$G = [I - P + \boldsymbol{tu}^T]^{-1} + \boldsymbol{ef}^T + \boldsymbol{g\pi}^T.$$

(Hunter, 1982)

Parameters of G-inverses of I – P

If G is any g-inverse of $I-P$ with stat prob vector $\boldsymbol{\pi}^T$ then G can be uniquely expressed in parametric form

$$G \equiv G(\boldsymbol{\alpha},\boldsymbol{\beta},\gamma) = [I-P+\boldsymbol{\alpha}\boldsymbol{\beta}^T]^{-1} + \gamma \mathbf{e}\boldsymbol{\pi}^T$$

where $\boldsymbol{\alpha},\boldsymbol{\beta}$ and γ involve $2m-1$ parameters with the properties $\boldsymbol{\pi}^T\boldsymbol{\alpha} = 1$ and $\boldsymbol{\beta}^T\mathbf{e} = 1$.

Given any G the parameters can be found as follows:

Let $A \equiv I-(I-P)G$ and $B = I-G(I-P)$ then

$A = \boldsymbol{\alpha}\boldsymbol{\pi}^T$ and $B = \mathbf{e}\boldsymbol{\beta}^T$ so that $\boldsymbol{\alpha} = A\mathbf{e}$ and $\boldsymbol{\beta}^T = \boldsymbol{\pi}^T B$.

Further $G\boldsymbol{\alpha} = (\gamma+1)\mathbf{e}$ and $\boldsymbol{\beta}^T G = (\gamma+1)\boldsymbol{\pi}^T$

with $\gamma+1 = \boldsymbol{\pi}^T G\boldsymbol{\alpha} = \boldsymbol{\beta}^T G\mathbf{e} = \boldsymbol{\beta}^T G\boldsymbol{\alpha}$. (Hunter, 1988)

$$G \in A\{1,\ 2\} \Leftrightarrow \gamma = -1,$$

$$G \in A\{1,\ 3\} \Leftrightarrow \boldsymbol{\alpha} = \boldsymbol{\pi},$$

$$G \in A\{1,\ 4\} \Leftrightarrow \boldsymbol{\beta} = \mathbf{e} / \mathbf{e}^T \mathbf{e} = \mathbf{e} / m,$$

$$G \in A\{1,\ 5a\} \Leftrightarrow \boldsymbol{\alpha} = \mathbf{e} \Leftrightarrow G\mathbf{e} = g\mathbf{e} \text{ for some } g,$$

$$G \in A\{1,\ 5b\} \Leftrightarrow \boldsymbol{\beta} = \boldsymbol{\pi} \Leftrightarrow \boldsymbol{\pi}^T G = h\boldsymbol{\pi}^T \text{ for some } h,$$

$$G \in A\{1,\ 5\} \Leftrightarrow \boldsymbol{\alpha} = \mathbf{e}, \boldsymbol{\beta} = \boldsymbol{\pi},$$

$$G \in A\{1,\ 2,\ 5\} \Leftrightarrow \boldsymbol{\alpha} = \mathbf{e}, \boldsymbol{\beta} = \boldsymbol{\pi},\ \gamma = -1.$$

Group inverse of $I - P$

The group inverse of $I - P$ has the form

$$A^{\#} = [I - P + \boldsymbol{e}\boldsymbol{\pi}^T]^{-1} - \boldsymbol{e}\boldsymbol{\pi}^T = [I - P + \Pi]^{-1} - \Pi \quad \text{(Meyer, 1975)}$$

Special properties:

(*i*) $(I - P)A^{\#} = I - \boldsymbol{e}\boldsymbol{\pi}^T$.

(*ii*) $A^{\#}(I - P) = I - \boldsymbol{e}\boldsymbol{\pi}^T$.

(*iii*) $A^{\#}e = \boldsymbol{0}$.

(*iv*) $\boldsymbol{\pi}^T A^{\#} = \boldsymbol{0}^T$.

Any matrix $A^{\#}$ satisfying $(i) - (iv)$ is the group inverse of $I - P$.

8. Solving for the stationary distribution

If $G = [I - P + \boldsymbol{t}\boldsymbol{u}^T]^{-1}$ where $\boldsymbol{u}$, $\boldsymbol{t}$ such that $\boldsymbol{u}^T\boldsymbol{e} \neq 0$, $\boldsymbol{\pi}^T\boldsymbol{t} \neq 0$,

$$\boldsymbol{\pi}^T = \frac{\boldsymbol{u}^T G}{\boldsymbol{u}^T G\boldsymbol{e}}.$$

(Paige,Styan,Wachter,1975), (Kemeny,1981), (Hunter,1982)

9. Solving for mean first passage times

(i) If G is any g-inverse of $I - P$, then

$$M = [G\Pi - E(G\Pi)_d + I - G + EG_d]D. \qquad \text{(Hunter, 1982)}$$

(ii) If $G\mathbf{e} = g\mathbf{e}$ for some g $\Leftrightarrow G \in A\{1,5a\}$

$$\Leftrightarrow \; M = [I - G + EG_d]D. \qquad \text{(Hunter, 2013)}$$

(iii) Let G be any g-inverse of $I - P$

then $H = G(I - \Pi)$ is a g-inverse of $I - P$ with $H\mathbf{e} = \mathbf{0}$

and $M = [I - H + EH_d]D$.

In particular, $M = [I - A^{\#} + EA_d^{\#}]D$, so that if $A^{\#} = [a_{ij}^{\#}]$,

$$m_{jj} = \frac{1}{\pi_j} \quad \text{and} \quad m_{ij} = \frac{a_{jj}^{\#} - a_{ij}^{\#}}{\pi_j}, \; (i \neq j).$$

10. Solving for the group inverse

$A^{\#}$ can be found from any g-inverse of $I - P$:

If G is any g-inverse of $I - P$, and $H = G(I - \Pi)$ then $K = (I - \Pi)H = (I - \Pi)G(I - \Pi)$ is a g-inverse of $I - P$ with $\boldsymbol{\pi}^T K = \mathbf{0}^T$ and $K\mathbf{e} = \mathbf{0}$.

In fact, $K = A^{\#}$, the group inverse of $I - P$.

10. Solving for the group inverse

Alternatively, $A^{\#}$ can be found from the m_{ij}:

Let $\tau_j \equiv \sum_{k=1}^{m} \pi_k m_{kj} = \sum_{k \neq j} \pi_k m_{kj} + 1,$

and let $A^{\#} = [a_{ij}^{\#}]$, then

$$a_{ij}^{\#} = \begin{cases} \pi_j(\tau_j - 1), & i = j, \\ \pi_j(\tau_j - 1 - m_{ij}) = a_{jj}^{\#} - \pi_j m_{ij}, & i \neq j. \end{cases}$$

(Ben-Ari, Neumann, 2012),(Hunter, 2013).

Computational considerations

Two relevant papers:
[1] (Heyman and O'Leary,1995) ("Computations with Markov chains" (2nd International Workshop on MC's)
[2] (Heyman and Reeves,1989) (ORSA J Computing)

[1]: "deriving means … of first passage times from … the group inverse $A^{\#}$ leads to a significant inaccuracy on the more difficult problems."

… "it does not make sense to compute …. the group generalized inverse unless the individual elements of those matrices are of interest."

Computational considerations - 2

From [2]: "The computation of M using $A^{\#}$ yields 3 sources of error:

1. The algorithm for computing $\boldsymbol{\pi}^T$
2. The computation of the inverse of $I - P + \Pi$

   (The matrix may have negative elements - can cause round-off errors in computing the inverse.)
3. The matrix evaluation of M

   (The matrix multiplying D may have negative elements). "

11. Perturbation procedures

The basic ideas are very simple:

Start with a transition matrix P_0, with known or easily evaluated stat prob vector $\boldsymbol{\pi}_0^T$, mean first passage time matrix M_0, and group inverse $A_0^\#$ for $I - P_0$, (or g-inverse G_0).

*S*equentially change $P_0 \to P_1 \to P_2 \to ...... \to P_m = P$ by replacing the i^{th} row of P_{i-1} with the i^{th} row, $\boldsymbol{p}_i^T$, of P (i = 1, 2, ..., m) to obtain P_i .

Thus, if $P_0 = \sum_{i=1}^{m} \boldsymbol{e}_i \boldsymbol{p}_{(0)i}^T$, and $P = \sum_{i=1}^{m} \boldsymbol{e}_i \boldsymbol{p}_i^T$

then $P_i = P_{i-1} + \boldsymbol{e}_i \boldsymbol{b}_i^T$ with $\boldsymbol{b}_i^T = \boldsymbol{p}_i^T - \boldsymbol{p}_{(0)i}^T$ for i = 1, 2, ..., m,

Update $\boldsymbol{\pi}_{i-1}^T$, M_{i-1} and $A_{i-1}^\#$ (or G_{i-1}) to $\boldsymbol{\pi}_i^T$, M_i and $A_i^\#$ (or G_i)

stopping with $\boldsymbol{\pi}_m^T = \boldsymbol{\pi}^T, M_i = M$ and $A_i^\# = A^\#$.

Choice of P_0

We require P_0 to be irreducible. The simplest structure is

$$P_0 = \begin{bmatrix} 1/m & 1/m & \ldots & 1/m & \ldots & 1/m \\ 1/m & 1/m & \ldots & 1/m & \ldots & 1/m \\ . & . & \ldots & . & . & . \\ 1/m & 1/m & \ldots & 1/m & \ldots & 1/m \\ . & . & \ldots & . & . & . \\ 1/m & 1/m & \ldots & 1/m & \ldots & 1/m \end{bmatrix} = \frac{1}{m}\mathbf{ee}^T = \frac{1}{m}E$$

This leads to

$$M_0 = m\mathbf{ee}^T \qquad = mE$$

and

$$\mathrm{A}_0^{\#} = I - \frac{1}{m}\mathbf{ee}^T = I - \frac{1}{m}E$$

12. The algorithms

We consider six algorithms:

1. Extend the procedure of Hunter (JAMSA, 1991) using a family of 1-condition generalized inverse updates to find successive stat prob vectors with extensions to the group inverses.
2. Consider successive direct row perturbation updates of the group inverse (and hence the mean first passage times).
3. Consider a blend of 1. and 2. through updating using matrix procedures for the stat probability vectors and the group inverses in tandem.

Algorithms 4, 5 and 6

Procedure based on updating specific g-inverses of I – P of the form $G = [I - P + \mathbf{e}\boldsymbol{\beta}^T]^{-1}$ that have simple forms for the mean first passage time matrix. since $G\mathbf{e} = g\mathbf{e}$.

4. $\boldsymbol{\beta}^T = \dfrac{\mathbf{e}^T}{m}$, with G = $G_e \equiv \left[I - P + \dfrac{\mathbf{e}\mathbf{e}^T}{m}\right]^{-1}$.
5. $\boldsymbol{\beta}^T = \mathbf{e}_1^T$, with G = $G_{e1} = [I - P + \mathbf{e}\mathbf{e}_1^T]^{-1}$
6. $\boldsymbol{\beta}^T = \mathbf{e}^T$, with G = $G_{ee} = [I - P + \mathbf{e}\mathbf{e}^T]^{-1}$

Note that it is easy to find the group inverse from the mean first passage time matrix since in these cases $A^{\#} = (I - \mathbf{e}\boldsymbol{\pi}^T)G$.

12.1: Perturbation procedures for stat distribns

With $P_0 = \boldsymbol{e}\boldsymbol{e}^T / m$,

$P_i = P_{i-1} + \mathbf{e}_i \boldsymbol{b}_i^T$ with $\boldsymbol{b}_i^T = \boldsymbol{p}_i^T - \mathbf{e}^T / m$.

With $\boldsymbol{t}_0 = \boldsymbol{e}$ and $\boldsymbol{u}_0^T = \mathbf{e}^T / m$ then $G_0 = [I - P_0 + \boldsymbol{t}_0 \boldsymbol{u}_0^T]^{-1} = I$.

Since $\boldsymbol{u}_0^T \boldsymbol{e} \neq 0$, $\boldsymbol{\pi}_0^T \boldsymbol{t}_0 \neq 0$, $\boldsymbol{\pi}_0^T = \dfrac{\boldsymbol{u}_0^T G_0}{\boldsymbol{u}_0^T G_0 \boldsymbol{e}} = \mathbf{e}^T / m$.

Let $\boldsymbol{t}_i = \mathbf{e}_i$ and $\boldsymbol{u}_i^T = \boldsymbol{u}_{i-1}^T + \boldsymbol{b}_i^T = \boldsymbol{u}_{i-1}^T + \boldsymbol{p}_i^T - \mathbf{e}^T / m$,

then $G_i = [I - P_i + \boldsymbol{t}_i \boldsymbol{u}_i]^{-1} = G_{i-1}[I + (\mathbf{e}_{i-1} - \mathbf{e}_i)(\boldsymbol{\pi}_{i-1}^T / \boldsymbol{\pi}_{i-1}^T \mathbf{e}_i)]$

implying $\boldsymbol{\pi}_i^T = \dfrac{\boldsymbol{u}_i^T G_i}{\boldsymbol{u}_i^T G_i \boldsymbol{e}}$, $i = 1, 2, ..., m$.

(Simplification of the calculations can be effected,

It can be shown that $G_i = G_{i-1} + F_{i-1}$ where all the

elements of $F_{i\text{-}1}$ in rows numbered i+1, ..., m, are zeros.)

Algorithm 1

(i) Let $G_0 = I$, $\boldsymbol{u}_0^T = \mathbf{e}^T / m$.

(ii) For $i = 1, 2, \ldots, m$, let $\boldsymbol{p}_i^T = \mathbf{e}_i^T P$,

$$\boldsymbol{u}_i^T = \boldsymbol{u}_{i-1}^T + \boldsymbol{p}_i^T - \mathbf{e}^T / m,$$

$$G_i = G_{i-1} + G_{i-1}(\mathbf{e}_{i-1} - \mathbf{e}_i)(\boldsymbol{u}_{i-1}^T G_{i-1} / \boldsymbol{u}_{i-1}^T G_{i-1} \mathbf{e}_i).$$

(iii) At $i = m$, let $G = G_m$ and

$$\boldsymbol{\pi}^T = \boldsymbol{\pi}_m^T = \frac{\boldsymbol{u}_m^T G_m}{\boldsymbol{u}_m^T G_m \mathbf{e}}.$$

(iv) Compute $H = G(I - \mathbf{e}\boldsymbol{\pi}^T)$.

(v) Compute $A^{\#} = (I - \mathbf{e}\boldsymbol{\pi}^T)H$.

(vi) Compute $M = [I - H + EH_d]D$ where $D = ((\mathbf{e}\boldsymbol{\pi}^T)_d)^{-1}$.

12.2: Perturbations of the Group Inverse

Let $\overline{P} = P + \mathrm{E}$ where the perturbing matrix E has the property $\mathrm{E}\boldsymbol{e} = \mathbf{0}$. Let $\Pi = \boldsymbol{e}\boldsymbol{\pi}^T$ where $\boldsymbol{\pi}^T$ is the stat prob vector of the MC associated with P.

Let $A^{\#}$ and $\overline{A}^{\#}$ be the group inverses of $A = I - P$ and $\overline{A} = I - \overline{P}$.

(i) $I - \mathrm{E}A^{\#}$ is non-singular,

(ii) the stat prob vector of the perturbed MC is

$$\overline{\pi}^T = \pi^T (I - \mathrm{E}A^{\#})^{-1}$$

(iii) the group inverse of $\overline{A} = I - \overline{P}$ is

$$\overline{A}^{\#} = A^{\#}(I - \mathrm{E}A^{\#})^{-1} - \Pi(I - \mathrm{E}A^{\#})^{-1}A^{\#}(I - \mathrm{E}A^{\#})^{-1}.$$

Row perturbations of the Group Inverse

Let $\mathrm{E} = \boldsymbol{e}_i \boldsymbol{b}^T$, i.e. a perturbation to the *i*-th row with $\boldsymbol{b}^T \mathbf{e} \neq 0$,

$$\bar{\boldsymbol{\pi}}^T = \boldsymbol{\pi}^T \left[I + \frac{1}{1 - \boldsymbol{b}^T A^\# \mathbf{e}_i} \boldsymbol{e}_i \boldsymbol{b}^T A^\# \right] \text{ and}$$

$$\bar{A}^\# = A^\# + \frac{1}{1 - \boldsymbol{b}^T A^\# \mathbf{e}_i} A^\# \mathbf{e}_i \boldsymbol{b}^T A^\# - \mathbf{e} \boldsymbol{y}^T,$$

$$\text{where} \quad \boldsymbol{y}^T = \left(\frac{\pi_i}{1 - \boldsymbol{b}^T A^\# \mathbf{e}_i} \right) \boldsymbol{b}^T \left(A^\# + \frac{\boldsymbol{b}^T (A^\#)^2 \mathbf{e}_i}{1 - \boldsymbol{b}^T A^\# \mathbf{e}_i} I \right) A^\#.$$

(Note that $\boldsymbol{y}^T \mathbf{e} = 0$.) See (Kirkland and Neumann, 2013).

Carry out row by row perturbations, with $\boldsymbol{b}_i^T$ the change at the *i*-th row, and $\mathrm{A}_i^\#$ the group inverse after the *i*-th change.

$$A_i^\# = R_i + \mathbf{e} \boldsymbol{y}_i^T \;\; \Rightarrow R_i = R_{i-1} + \frac{1}{1 - \boldsymbol{b}_i^T R_{i-1} \mathbf{e}_i} R_{i-1} \mathbf{e}_i \boldsymbol{b}_i^T R_{i-1} \text{ with } \boldsymbol{y}_i^T \mathbf{e} = 0.$$

Algorithm 2

(i) Let $P_0 = \mathbf{e}\,\mathbf{e}^T/m \Rightarrow A_0^\# = I - \mathbf{e}\,\mathbf{e}^T/m$. Take $R_0 = I - \mathbf{e}\,\mathbf{e}^T/m$.

(ii) For $i = 1, 2, \ldots, m$, let $\boldsymbol{p}_i^T = \mathbf{e}_i^T P$,

$$\boldsymbol{b}_i^T = \boldsymbol{p}_i^T - \mathbf{e}^T/m,$$

$$R_i = R_{i-1} + \frac{1}{1 - \boldsymbol{b}_i^T R_{i-1}\mathbf{e}_i} R_{i-1}\mathbf{e}_i \boldsymbol{b}_i^T R_{i-1}.$$

(iii) At $i = m$, let $R = R_m$ so that $A^\# = R + \mathbf{e}\boldsymbol{y}_m^T$.

$(I - P)A^\# = I - \mathbf{e}\boldsymbol{\pi}^T$ yields the stat prob vector:

$$\Rightarrow \boldsymbol{\pi}^T = \mathbf{e}_1^T - \mathbf{e}_1^T (I - P)R.$$

(iv) $\boldsymbol{\pi}^T A = \mathbf{0}^T$ *yields* the group inverse:

$$\Rightarrow \boldsymbol{y}_m^T = -\boldsymbol{\pi}^T R \Rightarrow A^\# = (I - \mathbf{e}\boldsymbol{\pi}^T)R.$$

(v) Compute $M = [I - A^\# + EA_d^\#]D$ where $D = ((\mathbf{e}\boldsymbol{\pi}^T)_d)^{-1}$.

12.3 Updating by matrix operations

Let $\overline{P} = P + \mathrm{E}$ where E has the property $\mathrm{E}\boldsymbol{e} = \mathbf{0}$.

Let $\Pi = \boldsymbol{e}\boldsymbol{\pi}^T$ and $\overline{\Pi} = \boldsymbol{e}\overline{\boldsymbol{\pi}}^T$ where $\boldsymbol{\pi}^T$ and $\overline{\boldsymbol{\pi}}^T$

are the stat prob vectors associated with P and $\overline{P}$.

$$\overline{\pi}^T = \pi^T (I - \mathrm{E}A^{\#})^{-1} \Rightarrow \overline{\Pi} = \Pi (I - \mathrm{E}A^{\#})^{-1}.$$

Under the perturbation $\mathrm{E} = \boldsymbol{e}_i \boldsymbol{b}^T$ to the i-th row with $\boldsymbol{b}^T\boldsymbol{e} \neq 0$,

$(I - \mathrm{E}A^{\#})^{-1} = I + \dfrac{1}{1 - \boldsymbol{b}^T A^{\#} \boldsymbol{e}_i} \boldsymbol{e}_i \boldsymbol{b}^T A^{\#}$ so that

$\overline{\Pi} = \Pi \left[I + \dfrac{1}{1 - \boldsymbol{b}^T A^{\#} \boldsymbol{e}_i} e_i \boldsymbol{b}^T A^{\#} \right]$ and

$$\overline{A}^{\#} = (I - \overline{\Pi}) A^{\#} (I - \mathrm{E}A^{\#})^{-1} = (I - \overline{\Pi}) A^{\#} \left(I + \frac{1}{1 - \boldsymbol{b}^T A^{\#} \boldsymbol{e}_i} \boldsymbol{e}_i \boldsymbol{b}^T A^{\#} \right).$$

Algorithm 3

(i) Let $P_0 = \boldsymbol{e}\boldsymbol{e}^T/m \Rightarrow \Pi_0 = \boldsymbol{e}\boldsymbol{e}^T/m,\ A_0^{\#} = I - \boldsymbol{e}\boldsymbol{e}^T/m.$

(ii) For $i = 1, 2, \ldots, m$, let $\boldsymbol{p}_i^T = \boldsymbol{e}_i^T P,\ \boldsymbol{b}_i^T = \boldsymbol{p}_i^T - \boldsymbol{e}^T/m,$

$$\mathrm{S}_i = I + \frac{1}{1 - \boldsymbol{b}_i^T A_{i-1}^{\#} \boldsymbol{e}_i} \boldsymbol{e}_i \boldsymbol{b}_i^T A_{i-1}^{\#},$$

$$\Pi_i = \Pi_{i-1} S_i,$$

$$\mathrm{A}_i^{\#} = (I - \Pi_i) A_{i-1}^{\#} S_i.$$

(iii) At $i = m$, let $\mathrm{S} = S_m$ then

$$\Pi = \Pi_{m-1} S,$$

$$\mathrm{A}^{\#} = (I - \Pi) A_{m-1}^{\#} S.$$

(iv) Compute $M = [I - A^{\#} + E A_d^{\#}] D$, where $D = (\Pi_d)^{-1}$.

12.4 Updating by g-inverses of I – P

From the Sherman-Morrison formula, with $P_0 = \dfrac{\mathbf{e}\mathbf{e}^T}{m}$

$$K_0 = [I - P_0 + \mathbf{e}\boldsymbol{\beta}^T]^{-1} = [I + \mathbf{e}\boldsymbol{h}^T]^{-1} = I - \frac{\mathbf{e}\boldsymbol{h}^T}{1 + \boldsymbol{h}^T\mathbf{e}}.$$

If $P_i = P_{i-1} + \mathbf{e}_i\boldsymbol{b}_i^T$, $K_i = [I - P_i + \mathbf{e}\boldsymbol{\beta}^T]^{-1} = K_{i-1} + \dfrac{1}{1 - \boldsymbol{b}_i^T\mathbf{e}_i} K_{i-1}\mathbf{e}_i\boldsymbol{b}_i^T K_{i-1}$.

4. $\boldsymbol{\beta}^T = \dfrac{\mathbf{e}^T}{m}, G_e = K_m,\ K_0 = I \Rightarrow \boldsymbol{\pi}^T = \dfrac{1}{m}\mathbf{e}^T K_m.$

5. $\boldsymbol{\beta}^T = \mathbf{e}_1^T, G_{e1} = K_m,\ K_0 = I + \mathbf{e}\left(\dfrac{\mathbf{e}^T}{m} - \mathbf{e}_1^T\right) \Rightarrow \boldsymbol{\pi}^T = \mathbf{e}_1^T K_m.$

6. $\boldsymbol{\beta}^T = \mathbf{e}^T, G_{ee} = K_m,\ K_0 = I - \left(\dfrac{m-1}{m}\right)\mathbf{e}\mathbf{e}^T \Rightarrow \boldsymbol{\pi}^T = \mathbf{e}^T K_m.$

Algorithm 4

(i) Let $K_0 = I$.

(ii) For $i = 1, 2, \ldots, m$, let $\boldsymbol{p}_i^T = \mathbf{e}_i^T P, \boldsymbol{b}_i^T = \boldsymbol{p}_i^T - \mathbf{e}^T / m$.

$$K_i = K_{i-1}(I + C_i) \text{ where } \mathrm{k}_i = 1 - \mathbf{e}_i K_{i-1} \mathbf{e}_i \text{ and } C_i = \frac{1}{k_i} \mathbf{e}_i \boldsymbol{b}_i^T K_{i-1}$$

(iii) At $i = m$, let $K = K_m$ then $\boldsymbol{\pi}^T = \dfrac{1}{m} \mathbf{e}^T K$.

(v) Compute $A^\# = (I - \mathbf{e}\boldsymbol{\pi}^T)K$.

(vi) Compute $M = [I - K + EK_d]D$ where $D = ((\mathbf{e}\boldsymbol{\pi}^T)_d)^{-1}$.

Algorithm 5

(i) Let $K_0 = I + \mathbf{e}\left(\frac{\mathbf{e}^T}{m} - \mathbf{e}_1^T\right)$

(ii) For $i = 1, 2, \ldots, m$, let $\boldsymbol{p}_i^T = \mathbf{e}_i^T P, \boldsymbol{b}_i^T = \boldsymbol{p}_i^T - \mathbf{e}^T / m$.

$$K_i = K_{i-1}(I + C_i) \text{ where } \mathrm{k}_i = 1 - \mathbf{e}_i K_{i-1} \mathbf{e}_i \text{ and } C_i = \frac{1}{k_i}\mathbf{e}_i \boldsymbol{b}_i^T K_{i-1}$$

(iii) At $i = m$, let $K = K_m$ then $\boldsymbol{\pi}^T = \mathbf{e}_1^T K$.

(v) Compute $A^{\#} = (I - \mathbf{e}\boldsymbol{\pi}^T)K$.

(vi) Compute $M = [I - K + EK_d]D$ where $D = ((\mathbf{e}\boldsymbol{\pi}^T)_d)^{-1}$.

Algorithm 6

(*i*) Let $K_0 = I - \left(\frac{m-1}{m}\right)\boldsymbol{ee}^T$.

(*ii*) For $i = 1, 2, ..., m$, let $\boldsymbol{p}_i^T = \mathbf{e}_i^T P, \boldsymbol{b}_i^T = \boldsymbol{p}_i^T - \mathbf{e}^T / m$.

$$K_i = K_{i-1}(I + C_i) \text{ where } \mathrm{k}_i = 1 - \mathbf{e}_i K_{i-1}\mathbf{e}_i \text{ and } C_i = \frac{1}{k_i}\mathbf{e}_i \boldsymbol{b}_i^T K_{i-1}$$

(*iii*) At $i = m$, let $K = K_m$ then $\boldsymbol{\pi}^T = \mathbf{e}^T K$.

(*v*) Compute $A^{\#} = (I - \mathbf{e}\boldsymbol{\pi}^T)K$.

(*vi*) Compute $M = [I - K + EK_d]D$ where $D = ((\mathbf{e}\boldsymbol{\pi}^T)_d)^{-1}$.

Test Problems

Introdced by Harrod &Plemmons (1984) and consided by others in different contexts.

TP1: The original transition matrix was not irreducible and replaced (Heyman (1987), Heyman & Reeves (1989)) by

$$\begin{bmatrix} .1 & .6 & 0 & .3 & 0 & 0 \\ .5 & .5 & 0 & 0 & 0 & 0 \\ .5 & .2 & 0 & 0 & .3 & 0 \\ 0 & .7 & 0 & .2 & 0 & .1 \\ .1 & 0 & .8 & 0 & 0 & .1 \\ .4 & 0 & .4 & 0 & 0 & .2 \end{bmatrix}$$

Test Problems

TP2 (Also Benzi (2004))

$$\begin{bmatrix} .85 & 0 & .149 & .0009 & 0 & .00005 & 0 & .00005 \\ .1 & .65 & .249 & 0 & .00009 & .00005 & 0 & .00005 \\ .1 & .8 & .09996 & .0003 & 0 & 0 & .0001 & 0 \\ 0 & .0004 & 0 & .7 & .2995 & 0 & .0001 & 0 \\ .0005 & 0 & .0004 & .399 & .6 & .0001 & 0 & 0 \\ 0 & .00005 & 0 & 0 & .00005 & .6 & .2499 & .15 \\ .00003 & 0 & .00003 & .00004 & 0 & .1 & .8 & .0999 \\ 0 & .00005 & 0 & 0 & .00005 & .1999 & .25 & .55 \end{bmatrix}$$

Test Problems

TP3

$$\begin{bmatrix} 0.999999 & 1.0\,E-07 & 2.0\,E-07 & 3.0\,E-07 & 4.0\,E-07 \\ 0.4 & 0.3 & 0 & 0 & 0.3 \\ 5.0\,E-07 & 0 & 0.999999 & 0 & 5.0\,E-07 \\ 5.0\,E-07 & 0 & 0 & 0.999999 & 5.0\,E-07 \\ 2.0\,E-07 & 3.0\,E-07 & 1.0\,E-07 & 4.0\,E-07 & 0.999999 \end{bmatrix}.$$

Test Problem TP41 $\equiv \varepsilon$ =1.0E-01, TP42 $\equiv \varepsilon$ =1.0E-03, TP43 $\equiv \varepsilon$ =1.0E-05, TP44 $\equiv \varepsilon$ =1.0E-07.

$$\begin{bmatrix}
.1-\varepsilon & .3 & .1 & .2 & .3 & \varepsilon & 0 & 0 & 0 & 0 \\
.2 & .1 & .1 & .2 & .4 & 0 & 0 & 0 & 0 & 0 \\
.1 & .2 & .2 & .4 & .1 & 0 & 0 & 0 & 0 & 0 \\
.4 & .2 & .1 & .2 & .1 & 0 & 0 & 0 & 0 & 0 \\
.6 & .3 & 0 & 0 & .1 & 0 & 0 & 0 & 0 & 0 \\
\varepsilon & 0 & 0 & 0 & 0 & .1-\varepsilon & .2 & .2 & .4 & .1 \\
0 & 0 & 0 & 0 & 0 & .2 & .2 & .1 & .3 & .2 \\
0 & 0 & 0 & 0 & 0 & .1 & .5 & 0 & .2 & .2 \\
0 & 0 & 0 & 0 & 0 & .5 & .2 & .1 & 0 & .2 \\
0 & 0 & 0 & 0 & 0 & .1 & .2 & .2 & .3 & .2
\end{bmatrix}$$

13. Computational Comparisons- Stat Distrns

Used the GTH algorithm to get accurate results for the stationary probabilities.

Used expressions for the exact results (when available)

Comparisons using MatLab with the 6 algorithms in single and double precision.

Expressions for

$\{\pi_i(E)\}, \{\pi_i(S)\}, \{\pi_i(D)\}, \{\pi_i(GTHS)\}, \{\pi_i(GTHD)\},$

Various comparisons between procedures A & B:

$$RE(A,B) = \sum_{i=1}^{m} \left|\pi_i(A) - \pi_i(B)\right|$$

$$MAXE(A,\ B) = \max_{1 \le i \le m} \left|\pi_i(A) - \pi_i(B)\right| \quad \text{(or MINE(A,B))}$$

$$MAXRE(A) = \max_{1 \le j \le m} \left|\pi_j(A) - \sum_{i=1}^{m} \pi_i(A) p_{ij}\right|$$

13. Computational Comparisons- MFPT

No expressions for the exact results are available.

Comparisons using MatLab with the 6 algorithms in single and double precision:

Expressions for $M(S)=[m_{ij}(S)],\ \ M(D)=[m_{ij}(D)]$

Various comparisons

$$RE\ M(S,\,D)=\sum_{i=1}^{m}\sum_{j=1}^{m}\left|m_{ij}(S)-m_{ij}(D)\right|,$$

$$MAXE\ M(S,\,D)=\max_{1\le i\le m,1\le j\le m}\left|m_{ij}(S)-m_{ij}(D)\right|,$$

$$MINE\ M(S,\,D)=\min_{1\le i\le m,1\le j\le m}\left|m_{ij}(S)-m_{ij}(D)\right|$$

$$MAXRES\ M(D)=\max_{1\le i\le m,1\le j\le m}\left|m_{ij}(D)-\sum_{k\ne j}p_{ik}m_{kj}(D)-1\right|.$$

13. Computational Comparisons- Group Inverse

No exact results available.

Comparisons difficult due to a number of conditions to be satisfied.

Comparisons using MatLab with the 6 algorithms in single and double precision for errors incurred in calculating the parameters.

$\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and γ, which should be close to

e, $\boldsymbol{\pi}$ and -1, respectively.

Compute MAXDELTA $\boldsymbol{\alpha} = \max_{1 \le i \le m} \left| \alpha_i - 1 \right|$,

$$\text{MAXDELTA } \boldsymbol{\beta} = \max_{1 \le i \le m} \left| \beta_i - \pi_i \right|,$$

$$\text{DELTA } \gamma = \left| \boldsymbol{\beta} A^{\#} \boldsymbol{\alpha} \right|$$

Comparisons

For this talk we present comparisons for the 7 test problems, the 6 algorithms under double precision for the

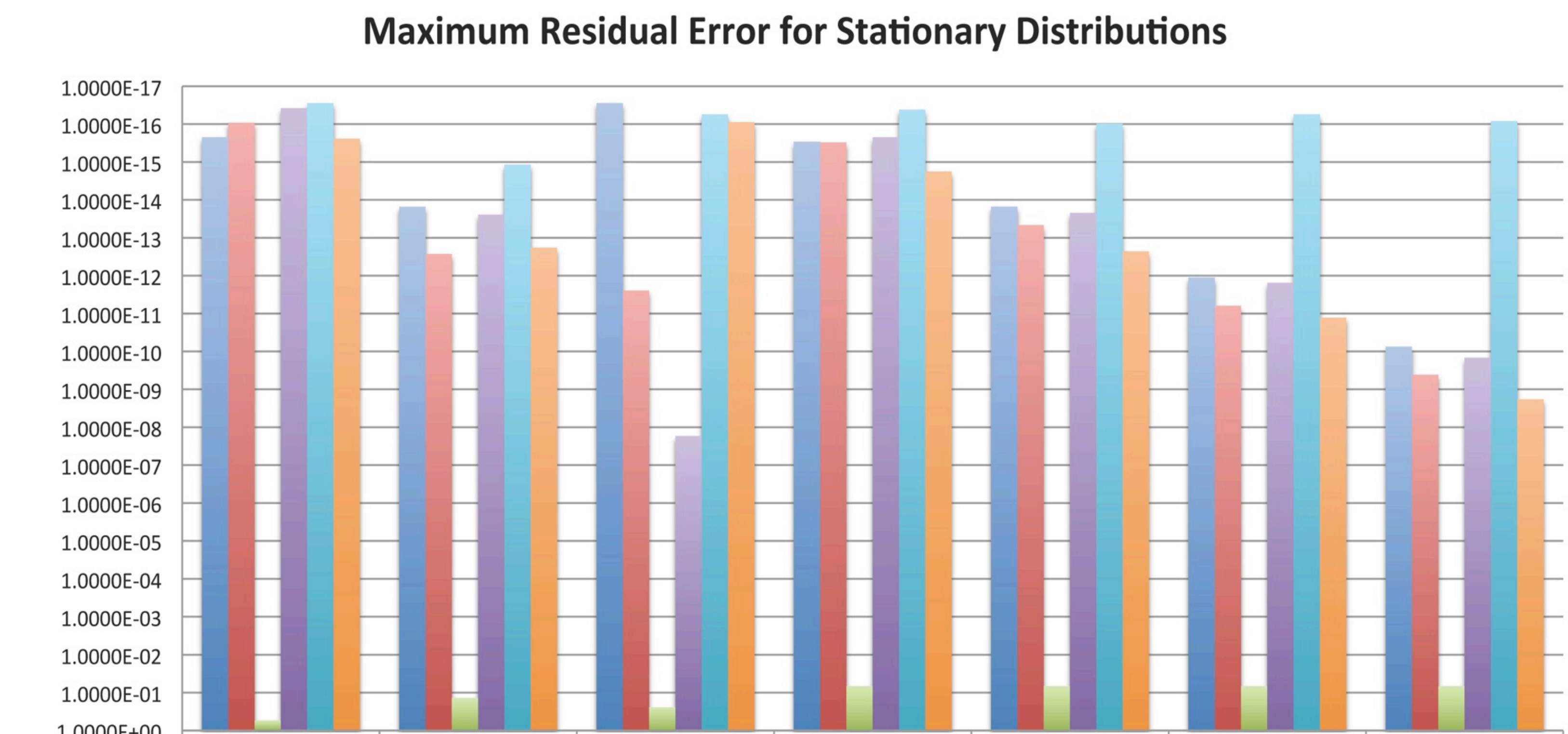
MAX RESIDUAL ERRORS for the stat probs,

$$\text{MAXRE SD} = \max_{1\le j\le m}\left|\pi_j - \sum_{i=1}^{m}\pi_i p_{ij}\right|,$$

MAX RESIDUAL ERRORS for the MFPTs

$$\text{MAXRES } M = \max_{1\le i\le m, 1\le j\le m}\left|m_{ij} - \sum_{k\ne j} p_{ik}m_{kj} - 1\right|$$

MAX DELTAS for the group inverse parameters

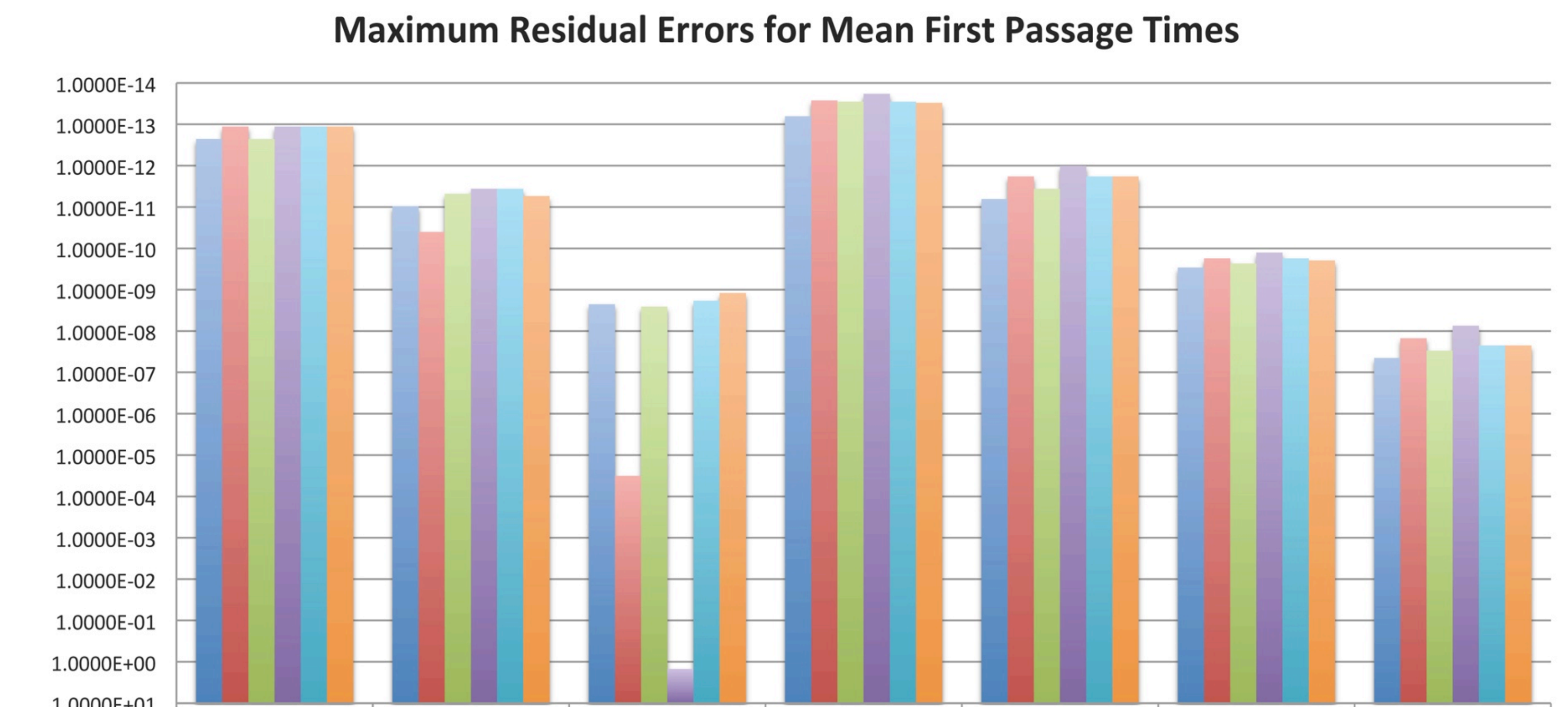


	TP1	TP2	TP3	TP41	TP42	TP43	TP44
Algorithm 1	2.2204E-16	1.5203E-14	2.7756E-17	2.9143E-16	1.5030E-14	1.1219E-12	7.4150E-11
Algorithm 2	9.1940E-17	2.6626E-13	2.4663E-12	3.0531E-16	4.6227E-14	6.1972E-12	4.1036E-10
Algorithm 3	5.4381E-01	1.3944E-01	2.4601E-01	6.7580E-02	6.7580E-02	6.7580E-02	6.7580E-02
Algorithm 4	3.8164E-17	2.4550E-14	1.7099E-08	2.2204E-16	2.2052E-14	1.5462E-12	1.4552E-10
Algorithm 5	2.7756E-17	1.1657E-15	5.5511E-17	4.1633E-17	9.7145E-17	5.5511E-17	8.3267E-17
Algorithm 6	2.4286E-16	1.8224E-13	8.8774E-17	1.7764E-15	2.2737E-13	1.2733E-11	1.8161E-09

AL3 performs the worst

AL5 consistently performs the best

All other algorithms perform satisfactorily

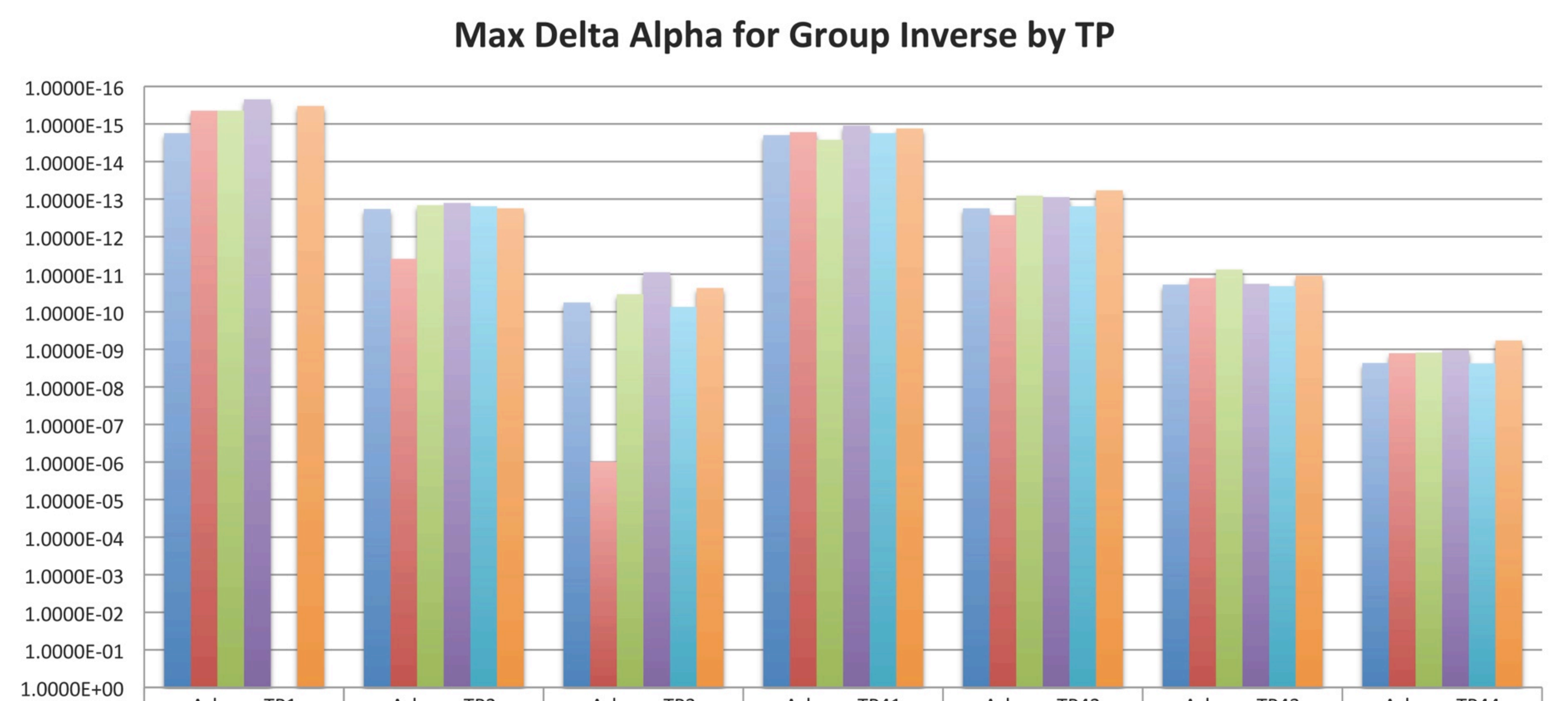


	TP1	TP2	TP3	TP41	TP42	TP43	TP44
Algorithm 1	2.2737E-13	9.5497E-12	2.2311E-09	6.3949E-14	6.4366E-12	2.9104E-10	4.5013E-08
Algorithm 2	1.1369E-13	3.9938E-11	3.1684E-05	2.6645E-14	1.8190E-12	1.7462E-10	1.4901E-08
Algorithm 3	2.2737E-13	4.7893E-12	2.5611E-09	2.8422E-14	3.6380E-12	2.3283E-10	2.9802E-08
Algorithm 4	1.1369E-13	3.6380E-12	1.4870E+00	1.8208E-14	1.0183E-12	1.2720E-10	7.4506E-09
Algorithm 5	1.1369E-13	3.6380E-12	1.8627E-09	2.8422E-14	1.8190E-12	1.7462E-10	2.2352E-08
Algorithm 6	1.1369E-13	5.4570E-12	1.2102E-09	3.0198E-14	1.8190E-12	1.9570E-10	2.2352E-08

All algorithms have a similar performance

Except AL4 and AL2 poor for TP3 and AL2 poor for TP2

AL4 best performer for other TPs

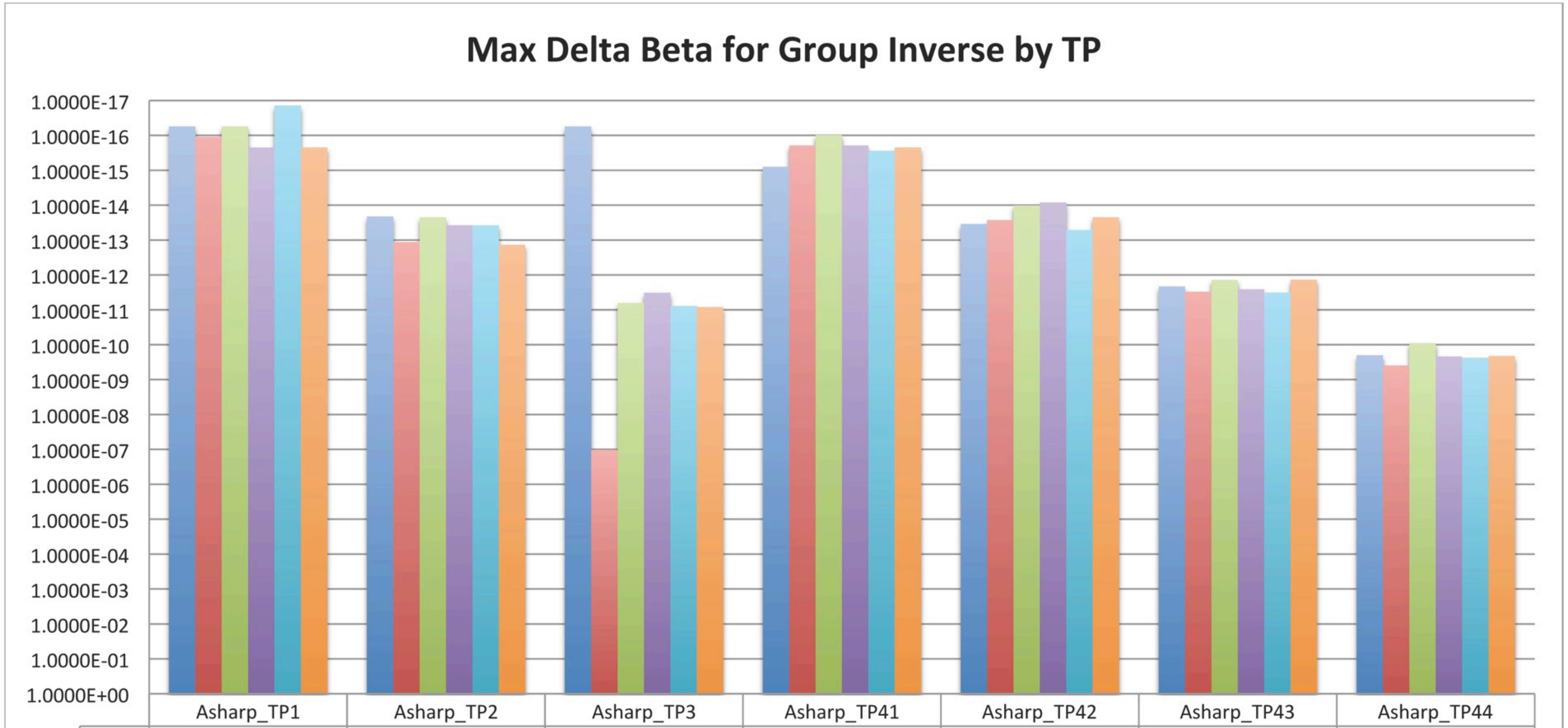


	Asharp_TP1	Asharp_TP2	Asharp_TP3	Asharp_TP41	Asharp_TP42	Asharp_TP43	Asharp_TP44
AL 1	1.7764E-15	1.8474E-13	5.6572E-11	1.9984E-15	1.7764E-13	1.9099E-11	2.3283E-09
AL 2	4.4409E-16	3.9182E-12	9.8175E-07	1.6653E-15	2.7001E-13	1.2733E-11	1.2806E-09
AL 3	4.4409E-16	1.4622E-13	3.4377E-11	2.6645E-15	8.1490E-14	7.5033E-12	1.2224E-09
AL 4	2.2204E-16	1.2790E-13	8.9113E-12	1.1102E-15	8.8818E-14	1.8190E-11	1.0477E-09
AL 5	0.0000E+00	1.5632E-13	7.4395E-11	1.7764E-15	1.5543E-13	2.0918E-11	2.3865E-09
AL 6	3.3307E-16	1.7764E-13	2.3463E-11	1.3323E-15	5.8620E-14	1.0914E-11	5.8208E-10

All algorithms have a similar performance

Except AL4 and AL2 poor for TP3 and AL2 poor for TP2

AL4 best performer for other TPs

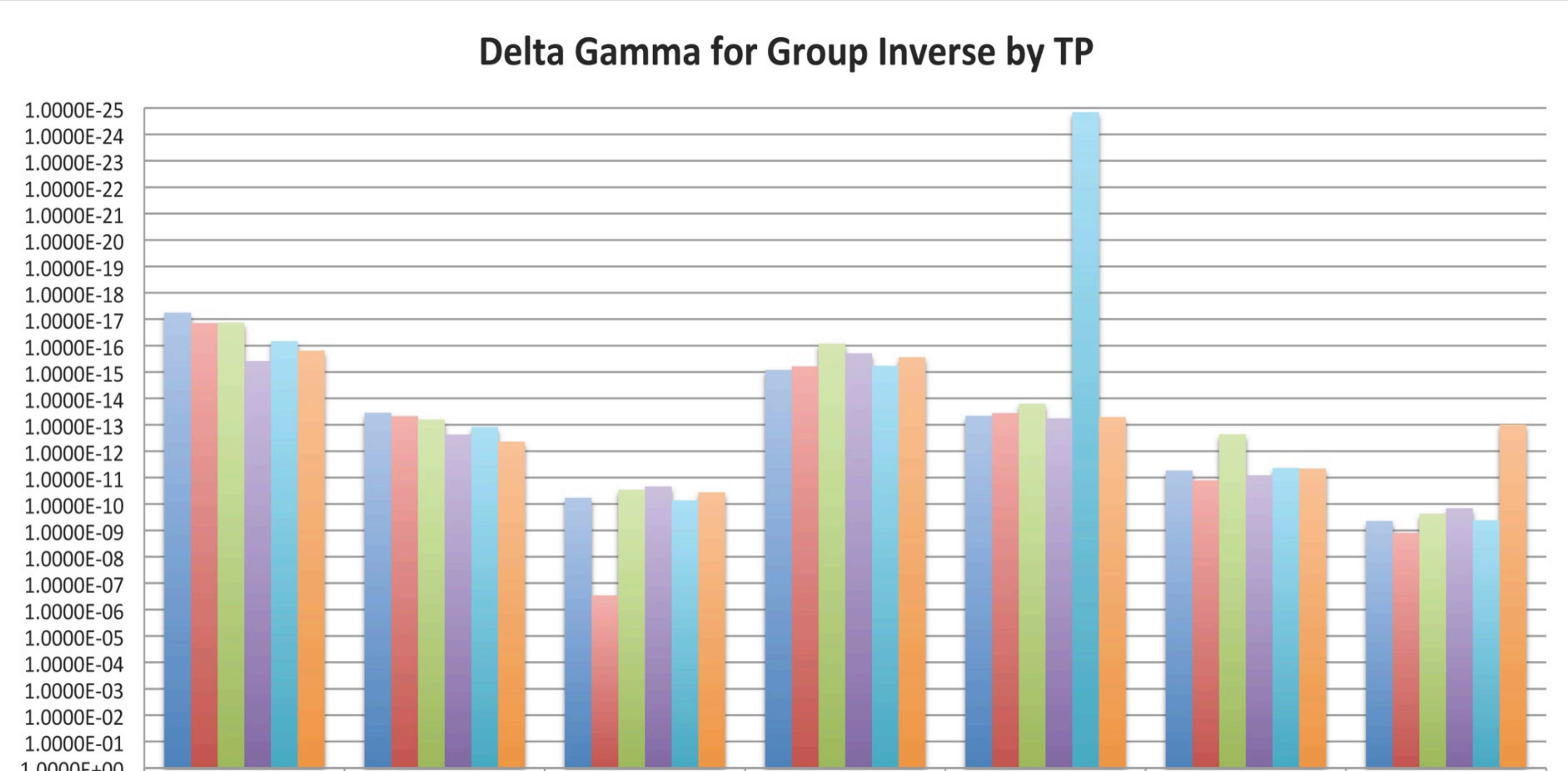


	Asharp_TP1	Asharp_TP2	Asharp_TP3	Asharp_TP41	Asharp_TP42	Asharp_TP43	Asharp_TP44
AL 1	5.5511E-17	2.1205E-14	5.5511E-17	7.9103E-16	3.4694E-14	2.1411E-12	1.9926E-10
AL 2	1.1102E-16	1.1385E-13	1.0386E-07	1.9429E-16	2.6867E-14	3.0336E-12	3.9361E-10
AL 3	5.5511E-17	2.2357E-14	6.3490E-12	9.7145E-17	1.0159E-14	1.3945E-12	8.9560E-11
AL 4	2.2204E-16	3.7678E-14	3.2374E-12	1.9429E-16	8.3683E-15	2.5772E-12	2.1732E-10
AL 5	1.3878E-17	3.7942E-14	7.7535E-12	2.7756E-16	5.1514E-14	3.2103E-12	2.3672E-10
AL 6	2.2204E-16	1.3556E-13	8.4040E-12	2.2204E-16	2.2204E-14	1.3702E-12	2.1036E-10

All algorithms have a similar performance

Except AL2 poor for TP3

AL1 very accurate for TP3



	Asharp_TP1	Asharp_TP2	Asharp_TP3	Asharp_TP41	Asharp_TP42	Asharp_TP43	Asharp_TP44
AL 1	5.6379E-18	3.5527E-14	5.8208E-11	8.3267E-16	4.6185E-14	5.4570E-12	4.3662E-10
AL 2	1.4068E-17	4.6962E-14	2.9238E-07	6.1062E-16	3.6193E-14	1.2695E-11	1.2224E-09
AL 3	1.3444E-17	6.3949E-14	2.9104E-11	8.3267E-17	1.5987E-14	2.2737E-13	2.3284E-10
AL 4	3.8207E-16	2.3448E-13	2.1828E-11	1.9429E-16	5.6843E-14	8.1855E-12	1.4596E-10
AL 5	6.7654E-17	1.2079E-13	7.2760E-11	5.8287E-16	1.4328E-25	4.3201E-12	4.0720E-10
AL 6	1.5396E-16	4.3343E-13	3.6380E-11	2.7756E-16	5.1514E-14	4.5475E-12	9.8060E-14

All algorithms have a similar performance

Except AL2 poor for TP3

AL5 very accurate for TP42