Convergence Issues of Iterative Aggregation/Disaggregation

Ivo Marek

Petr Mayer

Czech Institute of Technology, School of Civil Engineering, Thakurova 7,
166 29 Praha 6, Czech Republic

August 20, 2007 Computational Algebra Harrachov

Outline

1. Some motivation
2. IAD method for stationary probability vector
3. IAD with right hand side
4. Error formula
5. Fast konvergence
6. Conclusion

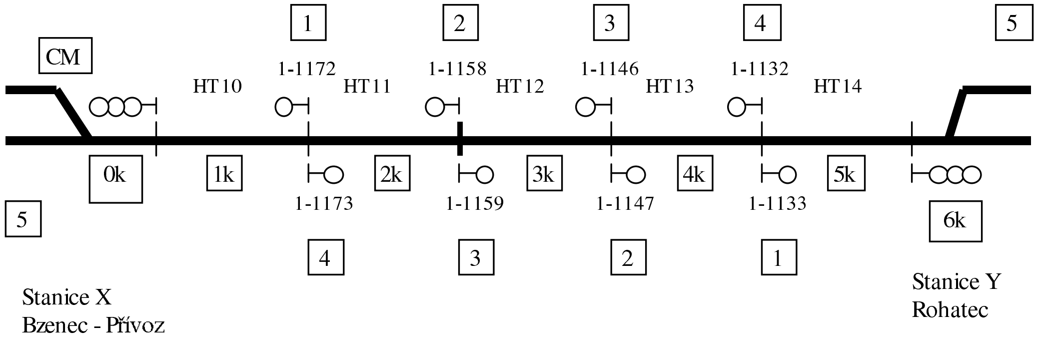
Some motivation

Definition 1. *Let elements of $T \in \Re^{n \times n}$ be non negative and $Te = e$, where $e = (1, ..., 1)^T \in \Re^n$. Then we call T the stochastic matrix.*

Definition 2. *A finite Markov chain is stochastic process, which moves through finite number of states, and for which the probability of entering a certain state depends only on the last state occupied.*

Definition 3. *A transient state has a non-zero probability that the chain will never return to this state.*
A reccurent(persistent) state has a zero probability that the chain will never return to this state.

E ... matrix of all ones
e ... vector of all ones
I ... identity matrix


1
2
3
4
5
CM
1-1172
1-1158
1-1146
1-1132
HT10
HT11
HT12
HT13
HT14
0k
1k
2k
3k
4k
5k
6k
1-1173
1-1159
1-1147
1-1133
5
4
3
2
1
Stanice X
Bzenec - Přívoz
Stanice Y
Rohatec

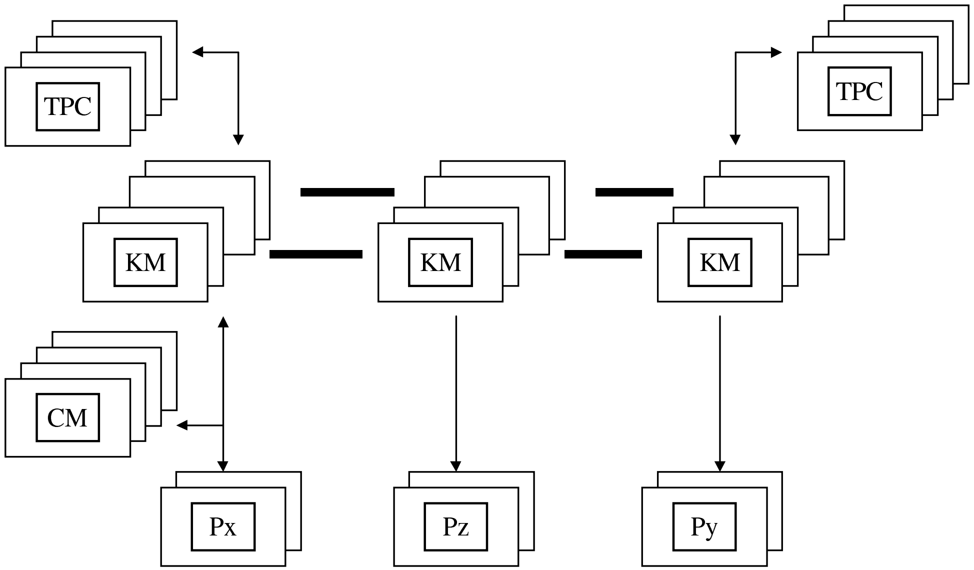

TPC
TPC
KM
KM
KM
CM
Px
Pz
Py

Figure 1.

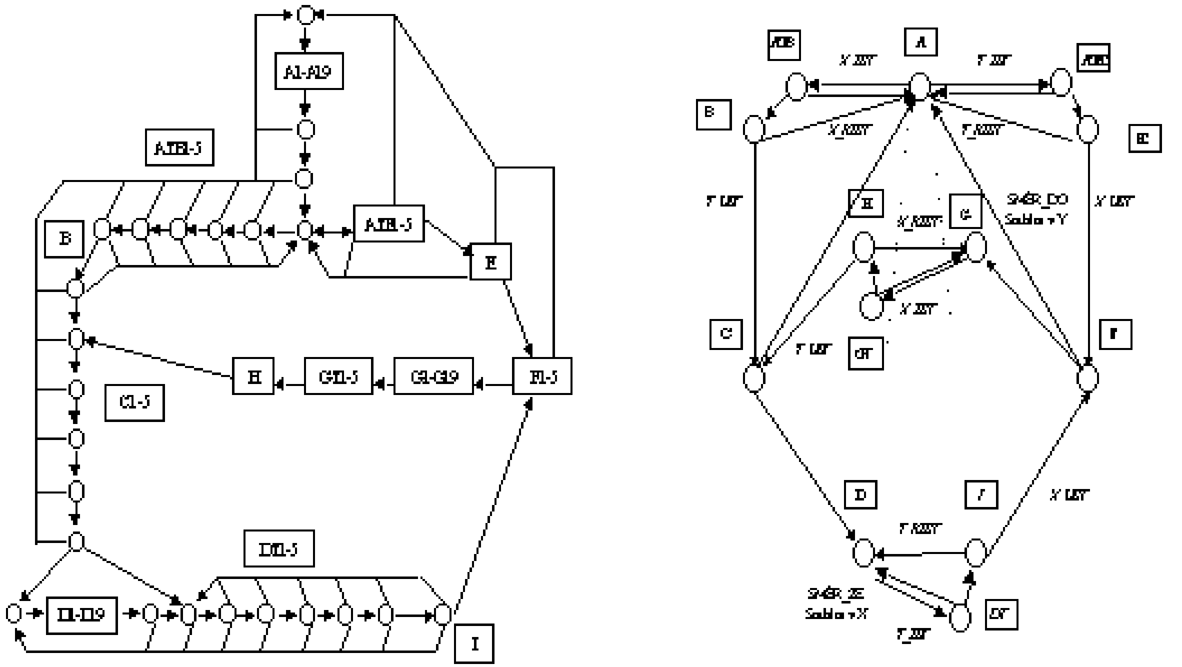


Figure 2.

To find long time behaviour of such system, we have to solve

Problem 1. We solve

$$Tx = x,\; e^T x = 1 \tag{1}$$

Let $Q \in \Re^{N \times N}$, such that $e^T Q = 0$, $\operatorname{diag} Q \leqslant 0$, $\operatorname{offdiag} Q \geqslant 0$, we try to compute

$$u(t) = e^{Qt} u(0)$$

using the Implicit Euler method, then we have to solve at every step system

$$u(t) = \tau Q\, u(t) + u(t - \tau).$$

After some rearangement ve finish with system of the type

$$x = Tx + b, \tag{2}$$

where T is nonnegative matrix with spectral radius less than one.

IAD method for stationary probability vector

$$g\colon \{1, \ldots, N\} \longrightarrow \{1, \ldots, n\}.$$

The restriction matrix $R \in \Re^{N \times n}$:

$$R_{g(i),i} = 1$$

$$(Rx)_j = \sum_{j=1,\, g(j)=i}^{N} x_j.$$

The prolongation matrix $S(x)$ is parametrised by vector $x \in \Re^N$, the nonzero elements of this matrix are

$$(S(x))_{i,g(i)} = \frac{x_i}{(Rx)_{g(i)}},$$

$$(S(x)z)_i = z_{g(i)} \frac{x_i}{(Rx)_{g(i)}} .$$

Aggregated matrix : $A(x) = RTS(x)$.

Lemma 4. *Let T be a column stochastic matrix, let g be an aggregation mapping and $x \in \Re^N$ such that $x \geqslant 0$ and $R\,x > 0$. Then aggregated matrix $A(x)$ is collumn stochastic. If the matrix T is irreducible and the vector x is strictly positive, then $A(x)$ is irreducible.*

Note 5. Let us note that the strict positivity of x is essential.

$$T = \begin{pmatrix} \frac{1}{3} & \frac{1}{4} & 0 & \frac{1}{4} \\ \frac{2}{3} & \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} & \frac{1}{4} \end{pmatrix}, \quad x = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{pmatrix}, \quad g\colon \begin{array}{c} 1 \mapsto 1 \\ 2 \mapsto 1 \\ 3 \mapsto 2 \\ 4 \mapsto 2 \end{array}.$$

We get the matrix $A(x) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, which is reducible.

Algorithm IAD(input: $T, M, W, x_{\text{init}}, \varepsilon, g, s$ output: x)

1. $k := 1, x_1 := x_{\text{init}}$
2. while $\|Tx_k - x_k\| > \varepsilon$ do
3. $\tilde{x} := (M^{-1}W)^s x_k$
4. $A(\tilde{x}) := RTS(\tilde{x})$
5. solve $A(\tilde{x})z = z$ and $e^T z = 1$
6. $k := k + 1$
7. $x_k = S(\tilde{x})z$
8. end while

Convergence theory for IAD can be found in [1].

Theorem 6. *Let T be a column stochastic matrix, let $\hat{x}$ be the solution of (1), then there exist s_0 and neighborhood of $\hat{x}$ such that for any x_{init} from this neighborhood and any $s > s_0$ the Algoritm IAD is convergent.*

IAD with right hand side

we solve problem (2), i.e.

$$x = Tx + b$$

Algorithm RHS(input: $T, M, W, x_{\text{init}}, \varepsilon, g, s$ output: x)

1. $k := 1, x_1 := x_{\text{init}}$
2. while $\|Tx_k - x_k\| > \varepsilon$ do
3. $\tilde{x}_0 = x_k$
4. for j=1,s do
5. $\tilde{x}_j := (M^{-1}W)\tilde{x}_{j-1} + M^{-1}b$
6. end do
7. $\tilde{x} = \tilde{x}_s$
8. $A(\tilde{x}) := RTS(\tilde{x})$
9. solve $z = A(\tilde{x})z + Rb$
10. $k := k + 1$
11. $x_k = S(\tilde{x})z$
12. end while

Error formula

For both previous processes we have same error formula

$$x_k - x^* = (M^{-1}W)^s (I - P(x_{k-1})T)^{-1}(I - P(x_{k-1}))$$

where

$$P(x) = S(x)\,R$$

Fast convergence

Theorem 7. *Let for splitting* M, W *be* $\mathrm{range}(M^{-1}W) \subseteq \mathrm{range}(S(\hat{x}))$. *Then Algorithm IAD terminates after the first iteration.*

Example 8. Let

$$T = \begin{pmatrix} 0.1 & 0.1 & 0.1 & 0.05 & 0.15 & 0.25 \\ 0.5 & 0.2 & 0 & 0.02 & 0.06 & 0.10 \\ 0.2 & 0.1 & 0.1 & 0.03 & 0.09 & 0.15 \\ 0.04 & 0.12 & 0.16 & 0.2 & 0.2 & 0.1 \\ 0.08 & 0.24 & 0.32 & 0.6 & 0.2 & 0.1 \\ 0.08 & 0.24 & 0.32 & 0.1 & 0.3 & 0.3 \end{pmatrix}$$

and splitting $I - T = M - W$ be

$$M = \begin{pmatrix} 0.9 & -0.1 & -0.1 & 0 & 0 & 0 \\ -0.5 & 0.8 & 0 & 0 & 0 & 0 \\ -0.2 & -0.1 & 0.9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.8 & -0.2 & -0.1 \\ 0 & 0 & 0 & -0.6 & 0.8 & -0.1 \\ 0 & 0 & 0 & -0.1 & -0.3 & 0.7 \end{pmatrix}$$

$$W = \begin{pmatrix} 0 & 0 & 0 & 0.05 & 0.15 & 0.25 \\ 0 & 0 & 0 & 0.02 & 0.06 & 0.10 \\ 0 & 0 & 0 & 0.03 & 0.09 & 0.15 \\ 0.04 & 0.12 & 0.16 & 0 & 0 & 0 \\ 0.08 & 0.24 & 0.32 & 0 & 0 & 0 \\ 0.08 & 0.24 & 0.32 & 0 & 0 & 0 \end{pmatrix}$$

Let aggregation mapping be

$$g\colon \begin{aligned} 1 &\mapsto 1 \\ 2 &\mapsto 1 \\ 3 &\mapsto 1 \\ 4 &\mapsto 2 \\ 5 &\mapsto 2 \\ 6 &\mapsto 2 \end{aligned}$$

$$x_0 = (\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6})^T$$

$$\tilde{x} = (0.104124, 0.102577, 0.084536, 0.182906, 0.309402, 0.311111)^T$$

$$A(\tilde{x}) = \begin{pmatrix} 0.4849558 & 0.3319149 \\ 0.5150442 & 0.6680851 \end{pmatrix}$$

$$z = (0.3918901, 0.6081099)^T$$

$$x_1 = (0.140109, 0.138029, 0.113752, 0.138442, 0.234187, 0.235481)^T$$

It is not true for Algorithm RHS, we need to add condition

$$b \in \text{range}\,(S(\hat{x})).$$

Other possibility is to replace steps 8, 9, 11 by

Step 8 : $A(\tilde{x}_s - \tilde{x}_{s-1}) := RTS(\tilde{x}_s - \tilde{x}_{s-1})$
Step 9 : solve $z = A(\tilde{x})z + R(b - Tx_k)$
Step 11: $x_k = S(\tilde{x})z + x_{k-1}$

Algorithm RHS(input: $T, M, W, x_{\text{init}}, \varepsilon, g, s$ output: x)

1. $k := 1, x_1 := x_{\text{init}}$
2. while $\|Tx_k - x_k\| > \varepsilon$ do
3. $\tilde{x}_0 = x_k$
4. for j=1,s do
5. $\tilde{x}_j := (M^{-1}W)\tilde{x}_{j-1} + M^{-1}b$
6. end do
7. $\tilde{x} = \tilde{x}_s$
8. $A(\tilde{x}_s - \tilde{x}_{s-1}) := RTS(\tilde{x}_s - \tilde{x}_{s-1})$
9. solve $z = A(\tilde{x})z + R(b - Tx_k)$
10. $k := k + 1$
11. $x_k = S(\tilde{x})z + x_{k-1}$
12. end while

Remark

for every irreducible stochastic matrix T there exist

$$\lim_{k\to\infty} T^k e = x^* \text{ and } Tx^* = x^*$$

Nearly dyadic matrices: bl = 60 sz = 420

method	p=0 w=0.5	p=0 w=0.1	p=0 w=0.01	p=0.1 w=0.5	p=0.1 w=0.1	p=0.1 w=0.01	p=0.5 w=0.5	p=0.5 w=0.1	p=0.5 w=0.01
power	53	224	3130	44	142	1812	55	70	915
MM	16	25	27	12	23	27	12	19	27
Vant	1	1	1	11	10	9	12	12	10
KMS	1	1	1	9	9	7	9	10	8
Jacobi	55	71	76	54	71	77	59	75	82
G.S.	41	51	54	30	42	47	29	40	45

Conclusion

– IAD methods are best for computing of SPV

– for cyclic matrices the power method applied to all ones vector is a reasonable choice

– structure of solution is significant

– Algorithm RHS is applicable for computing moments of Markov chains

Bibliography

[1] I. Marek and P. Mayer. Iterative aggregation/disaggregation methods for computing some characteristic of of markov chains. In *Large Scale Scientific Computing*, pages 68–82, 2001. Third International Conference, LSSC 2001, Sozopol, Bulgaria.