

Universal algorithms for generalized discrete matrix Bellman equations with symmetric Toeplitz matrix*

Sergeï Sergeev[†]

February 2, 2008

Abstract

This paper presents two universal algorithms for generalized discrete matrix Bellman equations with symmetric Toeplitz matrix. The algorithms are semiring extensions of two well-known methods solving Toeplitz systems in the ordinary linear algebra.

1 Introduction

As observed by B.A. Carré[1, 2], the Gaussian elimination without pivoting can be viewed as a prototype for some algorithms on graphs. M. Gondran [3] and G. Rote [4] made this observation precise by proving that the Gaussian elimination, under certain conditions, can be applied to the linear systems of equations over semirings.

The notion of universal algorithm over semiring was introduced by G.L. Litvinov, V.P. Maslov and E.V. Maslova in [5, 6]. These papers are to be considered in the framework of publications [7, 8, 9, 10, 11, 12]) of the Russian idempotent school, and more generally, in

*Research supported by RFBR grant 05-01-00824 and joint RFBR/CNRS grant 05-01-02807.

[†]Department of Physics, Sub-Department of Quantum Statistics and Field Theory, M.V. Lomonosov Moscow State University, 119992 Leninskie Gory, GSP-2, Moscow, Russia. E-mail: sergiej@gmail.com

the framework of idempotent and tropical mathematics, see [13, 14, 15] and references therein. Essentially, an algorithm is called universal if it does not depend on the computer representation of data and on a specific realization of algebraic operations involved in the algorithm [6]. Linear algebraic universal algorithms include generalized bordering method, LU- and LDM-decompositions for solving matrix equations. These methods are basically due to B.A. Carré, see also [6].

It was observed in [5, 6] that universal algorithms can be implemented by means of objective-oriented programming supported by C++, MATLAB, Scilab, Maple and other computer systems and languages. Such universal programs can be instrumental in many areas including the problems of linear algebra, optimization theory, and interval analysis over positive semirings, see [5, 11, 12, 16].

This paper presents new universal algorithms based on the methods of Durbin and Levinson, see [17], Sect. 4.7. These algorithms solve systems of linear equations with symmetric Toeplitz matrices. Our universal algorithms have the same computational complexity $O(n^2)$ as their prototypes which beats the complexity $O(n^3)$ of the LDM-decomposition method. All algorithms are described as MATLAB-programs, meaning that they can be actually implemented.

The author is grateful to G.L. Litvinov and A.N. Sobolevskii for drawing his attention to this problem and for valuable discussions.

2 Semirings and universal algorithms

A set S equipped with *addition* \oplus and *multiplication* \odot is a *semiring (with unity)* if the following axioms hold:

- 1) (S, \oplus) is a commutative semigroup with neutral element $\mathbf{0}$;
- 2) (S, \odot) is a semigroup with neutral element $\mathbf{1} \neq \mathbf{0}$;
- 3) $a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$, $(a \oplus b) \odot c = (a \odot c) \oplus (b \odot c)$ for all $a, b, c \in S$ (distributivity);

4) $\mathbf{0} \odot a = a \odot \mathbf{0} = \mathbf{0}$ for all $a \in S$.

In the sequel, we omit the notation \odot whenever this is convenient and does not lead to confusion.

The semiring S is called *idempotent* if $a \oplus a = a$ for any $a \in S$. In this case \oplus induces the *canonical partial order* relation

$$a \preceq b \Leftrightarrow a \oplus b = b. \quad (1)$$

The semiring S is called *complete* (cf. [18]), if any subset $\{x_\alpha\} \subset S$ is summable and the infinite distributivity

$$\begin{aligned} c \odot (\bigoplus_\alpha x_\alpha) &= \bigoplus_\alpha (c \odot x_\alpha), \\ (\bigoplus_\alpha x_\alpha) \odot c &= \bigoplus_\alpha (x_\alpha \odot c). \end{aligned} \quad (2)$$

holds for all $c \in S$ and $\{x_\alpha\} \subset S$. This property is natural in idempotent semirings and also in the theory of partially ordered spaces (cf. G. Birkhoff [19]) with partial order (1). Complete idempotent semirings are called *a-complete* (cf. [9]).

Consider the *closure operation*

$$a^* = \bigoplus_{i=0}^{\infty} a^i. \quad (3)$$

In the complete semirings it is defined for all elements. The property

$$a^* = \mathbf{1} \oplus aa^* = \mathbf{1} \oplus a^*a, \quad (4)$$

reveals that the closure operation is a natural extension of $(1 - a)^{-1}$.

We give some examples of semirings living on the set of reals \mathbf{R} totally ordered by \leq : the semiring \mathbf{R}_+ with customary operations $\oplus = +$, $\odot = \cdot$ and neutral elements $\mathbf{0} = 0$ and $\mathbf{1} = 1$; the semiring $\mathbf{R}_{\max} = \mathbf{R} \cup \{-\infty\}$ with operations $\oplus = \max$, $\odot = +$, and neutral elements $\mathbf{0} = -\infty$, $\mathbf{1} = 0$; the semiring $\widehat{\mathbf{R}}_{\max} = \mathbf{R}_{\max} \cup \{\infty\}$, which is a completion of \mathbf{R}_{\max} with the element ∞ satisfying $a \oplus \infty = \infty$ for all a , $a \odot \infty = \infty \odot a = \infty$ for $a \neq \mathbf{0}$ and $\mathbf{0} \odot \infty = \infty \odot \mathbf{0} = \mathbf{0}$; the semiring $\mathbf{R}_{\max, \min} = \mathbf{R} \cup \{\infty\} \cup \{-\infty\}$ with $\oplus = \max$, $\odot = \min$, $\mathbf{0} = -\infty$, and $\mathbf{1} = \infty$.

Consider operation (3) for the examples above. In \mathbf{R}_+ the closure a^* equals $(1 - a)^{-1}$ if $a < 1$ and is undefined otherwise; in \mathbf{R}_{\max} it equals $\mathbf{1}$ if $a \leq \mathbf{1}$ and is undefined otherwise; in $\widehat{\mathbf{R}}_{\max}$ we have $a^* = \mathbf{1}$ for $a \leq \mathbf{1}$ and $a^* = \infty$ for $a > \mathbf{1}$; in $\mathbf{R}_{\max, \min}$ we have $a^* = \mathbf{1}$ for all a . Note that $\widehat{\mathbf{R}}_{\max}$ and \mathbf{R}_{\max} are a -complete, so the closure is defined for any element of these semirings.

The *matrix* operations \oplus and \odot are defined analogously to their counterparts in linear algebra. Denote by $\text{Mat}_{mn}(S)$ the set of all $m \times n$ matrices over the semiring S . By I_n we denote the $n \times n$ unity matrix, that is, the matrix with $\mathbf{1}$ on the diagonal and $\mathbf{0}$ off the diagonal. As usual, we have $AI_n = I_nA = A$ and $A^0 = I_n$ for any $A \in \text{Mat}_{nn}(S)$. The set $\text{Mat}_{nn}(S)$ of all $n \times n$ square matrices is a semiring. Its unity is I_n and its zero is $\mathbf{0}_n$, the square matrix with all entries equal to $\mathbf{0}$. If S is complete and/or idempotent, then so is the semiring $\text{Mat}_{nn}(S)$. If S (and hence $\text{Mat}_{nn}(S)$) is complete, the closure A^* is defined for any matrix A and it satisfies (4). Note that if S is partially ordered, then $\text{Mat}_{mn}(S)$ is ordered elementwise: $A \preceq B$ iff $A_{ij} \preceq B_{ij}$ for all $i = 1, \dots, m$ and $j = 1, \dots, n$. If S is idempotent and canonically ordered (1), then the elementwise order of $\text{Mat}_{mn}(S)$ also satisfies (1).

The closure operation of matrices is important for the (*discrete stationary*) *matrix Bellman equations*

$$X = AX \oplus B. \quad (5)$$

If the closure of A exists and (4) holds, then $X = A^*B$ is a solution to (5). In a -complete idempotent semirings the matrix A^*B is the least solution of this equation with respect to (1).

Since A^* is a generalization of $(I - A)^{-1}$, the known universal algorithms for A^* are generalizations of the methods for matrix inverses, and the known algorithms for Bellman equations are generalizations of the methods for $AX = B$. Further we consider the generalized bordering method.

Let A be a square matrix. Closures of its main submatrices A_k can be found inductively. The base of induction is A_1^* , the closure of the the first diagonal entry. Generally, we

represent A_{k+1} as

$$A_{k+1} = \begin{pmatrix} A_k & g_k \\ h_k^T & a_{k+1} \end{pmatrix},$$

assuming that we have found the closure of A_k . In this representation, g_k and h_k are columns with k entries and a_{k+1} is a scalar. We also represent A_{k+1}^* as

$$A_{k+1}^* = \begin{pmatrix} U_k & v_k \\ w_k^T & u_{k+1} \end{pmatrix}.$$

Using (4) we obtain that

$$\begin{aligned} u_{k+1} &= (h_k^T A_k^* g_k \oplus a_{k+1})^*, \\ v_k &= A_k^* g_k u_{k+1}, \\ w_k^T &= u_{k+1} h_k^T A_k^*, \\ U_k &= A_k^* g_k u_{k+1} h_k^T A_k^* \oplus A_k^*. \end{aligned} \tag{6}$$

Consider the bordering method for finding the solution $x = A^*b$ to (5), where $X = x$ and $B = b$ are column vectors. Firstly, we have $x^{(1)} = A_1^* b_1$. Let $x^{(k)}$ be the vector found after $(k-1)$ steps, and let us write

$$x^{(k+1)} = \begin{pmatrix} z \\ x_{k+1} \end{pmatrix}.$$

Using (6) we obtain that

$$\begin{aligned} x_{k+1} &= u_{k+1} (h_k^T x^{(k)} \oplus b_{k+1}), \\ z &= x^{(k)} \oplus A_k^* g_k x_{k+1}^{(k+1)}. \end{aligned} \tag{7}$$

We have to compute $A_k^* g_k$. In general it makes a problem, but not in the case of the next section when A is symmetrical Toeplitz.

We also note that the bordering method described by (6) and (7) is valid more generally over Conway semirings, see [18] for the definition.

3 Universal algorithms for Toeplitz linear systems

Formally, a matrix $A \in \text{Mat}_{nn}(S)$ is called (*generalized*) *Toeplitz* if there exist scalars $r_{-n+1}, \dots, r_0, \dots, r_{n-1}$ such that $A_{ij} = r_{j-i}$ for all i and j . Informally, Toeplitz matrices are such that their entries are constant along any line parallel to the main diagonal (and along the main diagonal itself). For example,

$$A = \begin{pmatrix} r_0 & r_1 & r_2 & r_4 \\ r_{-1} & r_0 & r_1 & r_2 \\ r_{-2} & r_{-1} & r_0 & r_1 \\ r_{-3} & r_{-2} & r_{-1} & r_0 \end{pmatrix} \quad (8)$$

is Toeplitz. Such matrices are not necessarily symmetric. However, they are always *persymmetric*, that is, symmetric with respect to the inverse diagonal. This property is algebraically expressed as $A = E_n A^T E_n$, where $E_n = [e_n, \dots, e_1]$. By e_i we denote the column whose i th entry is $\mathbf{1}$ and other entries are $\mathbf{0}$. The property $E_n^2 = I_n$ (where I_n is the $n \times n$ identity matrix) implies that the product of two persymmetric matrices is persymmetric. Hence any degree of a persymmetric matrix is persymmetric, and so is the closure of a persymmetric matrix. Thus, if A is persymmetric, then

$$E_n A^* = (A^*)^T E_n. \quad (9)$$

Further we deal only with symmetric Toeplitz matrices. Consider the equation $y = T_n y \oplus r^{(n)}$, where $r^{(n)} = (r_1, \dots, r_n)^T$ and T_n is defined by the scalars r_0, r_1, \dots, r_{n-1} so that $T_{ij} = r_{|j-i|}$ for all i and j . This is a generalization of the Yule-Walker problem [17]. Assume that we have obtained a solution $y^{(k)}$ to the system $y = T_k y \oplus r^{(k)}$ for some k such that $1 \leq k \leq n-1$, where T_k is the main $k \times k$ submatrix of T_n . We write T_{k+1} as

$$T_{(k+1)} = \begin{pmatrix} T_k & E_k r^{(k)} \\ r^{(k)T} E_k & r_0 \end{pmatrix}.$$

We also write $y^{(k+1)}$ and $r^{(k+1)}$ as

$$y^{(k+1)} = \begin{pmatrix} z \\ \alpha_k \end{pmatrix}, \quad r^{(k+1)} = \begin{pmatrix} r^{(k)} \\ r_{k+1} \end{pmatrix}.$$

Using (7), (9) and the identity $T_k^* r^{(k)} = y^{(k)}$, we obtain that

$$\begin{aligned}\alpha_k &= (r_0 \oplus r^{(k)T} y^{(k)})^* (r^{(k)T} E_k y^{(k)} \oplus r_{k+1}), \\ z &= E_k y^{(k)} \alpha_k \oplus y^{(k)}.\end{aligned}\tag{10}$$

Denote $\beta_k = r_0 \oplus r^{(k)T} y^{(k)}$. The following argument shows that β_k can be found recursively if $(\beta_{k-1}^*)^{-1}$ exists.

$$\begin{aligned}\beta_k &= r_0 \oplus [r^{(k-1)T} \quad r_k] \begin{pmatrix} E_{k-1} y^{(k-1)} \alpha_{k-1} \oplus y^{(k-1)} \\ \alpha_{k-1} \end{pmatrix} = \\ &= r_0 \oplus r^{(k-1)T} y^{(k-1)} \oplus (r^{(k-1)T} E_{k-1} y^{(k-1)} \oplus r_k) \alpha_{k-1} = \\ &= \beta_{k-1} \oplus (\beta_{k-1}^*)^{-1} \odot (\alpha_{k-1})^2.\end{aligned}\tag{11}$$

The argument above is not always valid and this will make us write two versions of our algorithm, the first one involving (11) and the second one not involving it. We will write these two versions in one program and mark the expressions which refer only to the first or only to the second version by the MATLAB-style comments %1 or %2, respectively. Collecting the expressions for β_k, α_k and z , we obtain the following recursive expression for $y^{(k)}$:

$$\begin{aligned}\beta_k &= r_0 \oplus r^{(k)T} y^{(k)}, & \%2 \\ \beta_k &= \beta_{k-1} \oplus (\beta_{k-1}^*)^{-1} \odot (\alpha_{k-1})^2, & \%1 \\ \alpha_k &= (\beta_k)^* \odot ((r^{(k)T} E_k y^{(k)} \oplus r_{k+1}), \\ y^{(k+1)} &= \begin{pmatrix} E_k y^{(k)} \alpha_k \oplus y^{(k)} \\ \alpha_k \end{pmatrix}.\end{aligned}\tag{12}$$

Recursive expression (12) is a generalized version of the Durbin method for the Yule-Walker problem [17]. Using this expression we obtain the following algorithm.

Algorithm 1 *The Yule-Walker problem for the Bellman equations with symmetric Toeplitz matrix.*

```
function  $y = \text{durbin}(r_0, r)$ 
 $n = \text{size}(r) + 1$ 
 $y(1) = r_0^* \odot r(1)$ 
```

```

 $\beta = r_0$  %1
 $\alpha = r_0^* \odot r(1)$  %1
for  $k = 1 : n - 1$ 
 $\beta = r_0 \oplus r(1 : k) \odot y(1 : k)$  %2
 $\beta = \beta \oplus (\beta^*)^{-1} \odot \alpha^2$  %1
 $\alpha = \beta^* \odot (r(k : -1 : 1) \odot y(1 : k) \oplus r(k + 1))$ 
 $z(1 : k) = y(1 : k) \oplus \alpha \odot y(k : -1 : 1)$ 
 $y(1 : k) = z(1 : k)$ 
 $y(k + 1) = \alpha$ 
end

```

Now we consider the problem of finding $x^{(n)} = T_n^* b^{(n)}$ where T_n is as above and $b^{(n)} = (b_1, \dots, b_n)^T$ is arbitrary. We also introduce the column $y^{(n)}$ which solves the Yule-Walker problem: $y^{(n)} = T_n^* r^{(n)}$. The main idea is to find the expression for $x^{(k+1)} = T_{k+1}^* b^{(k+1)}$ involving $x^{(k)}$ and $y^{(k)}$. We write $x^{(k+1)}$ and $b^{(k+1)}$ as

$$x^{(k+1)} = \begin{pmatrix} v \\ \mu_k \end{pmatrix}, \quad b^{(k+1)} = \begin{pmatrix} b^{(k)} \\ b_{k+1} \end{pmatrix}.$$

Making use of the persymmetry of T_k^* and of the identities $T_k^* b_k = x^{(k)}$ and $T_k^* r_k = y^{(k)}$, we specialize expressions (7) and obtain that

$$\begin{aligned} \mu_k &= (r_0 \oplus r^{(k)T} y^{(k)})^* \odot ((r^{(k)T} E_k x^{(k)} \oplus b_{k+1}), \\ v &= E_k y^{(k)} \mu_k \oplus x^{(k)}. \end{aligned} \tag{13}$$

The coefficient $r_0 \oplus r^{(k)T} y^{(k)} = \beta_k$ is again to be expressed as $\beta_k = \beta_{k-1} \oplus (\beta_{k-1}^*)^{-1} \odot (\alpha_{k-1})^2$, if the closure $(\beta_{k-1})^*$ is invertible. Using this we obtain the following recursive expression:

$$\begin{aligned} \beta_k &= r_0 \oplus r^{(k)T} y^{(k)}, & \%2 \\ \beta_k &= \beta_{k-1} \oplus (\beta_{k-1}^*)^{-1} \odot (\alpha_{k-1})^2, & \%1 \\ \mu_k &= (\beta_k)^* \odot ((r^{(k)T} E_k x^{(k)} \oplus b_{k+1}), \\ x^{(k+1)} &= \begin{pmatrix} E_k y^{(k)} \mu_k \oplus x^{(k)} \\ \mu_k \end{pmatrix}. \end{aligned} \tag{14}$$

This expression yields the following generalized version of the Levinson algorithm for solving linear symmetric Toeplitz systems [17]:

Algorithm 2 *Bellman system with symmetric Toeplitz matrix.*

```

function  $y = \text{levinson}(r_0, r, b)$ 
 $n = \text{size}(b)$ 
 $y(1) = r_0^* \odot r(1); \quad x(1) = r_0^* \odot b(1)$ 
 $\beta = r_0 \quad \%1$ 
 $\alpha = r_0^* \odot r(1) \quad \%1$ 
for  $k = 1 : n - 1$ 
 $\beta = r_0 \oplus r(1 : k) \odot y(1 : k) \quad \%2$ 
 $\beta = \beta \oplus (\beta^*)^{-1} \odot \alpha^2 \quad \%1$ 
 $\mu = (r(k : -1 : 1) \odot x(1 : k) \oplus b(k + 1)) \odot \beta^*$ 
 $v(1 : k) = x(1 : k) \oplus \mu \odot y(k : -1 : 1)$ 
 $x(1 : k) = v(1 : k)$ 
 $x(k + 1) = \mu$ 
if  $k < n - 1$ 
 $\alpha = (r(k : -1 : 1) \odot y(1 : k) \oplus r(k + 1)) \odot \beta^*$ 
 $z(1 : k) = y(1 : k) \oplus \alpha \odot y(k : -1 : 1)$ 
 $y(1 : k) = z(1 : k)$ 
 $y(k + 1) = \alpha$ 
end
end

```

The computational complexity of all methods described in this section is $O(n^2)$.

References

- [1] B.A. Carré. An algebra for network routing problems. *J. of the Inst. of Maths. and Applics*, 7:273–299, 1971.

- [2] R.C. Backhouse and B.A. Carré. Regular algebra applied to path-finding problems. *J. of the Inst. of Maths. and Applics*, 15:161–186, 1975.
- [3] M. Gondran. Path algebra and algorithms. In B. Roy, editor, *Combinatorial programming: methods and applications*, pages 137–148. Reidel, Dordrecht, 1975.
- [4] G. Rote. A systolic array algorithm for the algebraic path problem. *Computing*, 34:191–219, 1985.
- [5] G.L. Litvinov and V.P. Maslov. Correspondence principle for idempotent calculus and some computer applications. In J. Gunawardena, editor, *Idempotency*, Publications of the I. Newton Institute, pages 420–443. Cambridge Univ. Press, 1998.
- [6] G. Litvinov and E. Maslova. Universal numerical algorithms and their software implementation. *Programming and Computer Software*, 26(5):275–380, 2000. E-print arXiv:math.NA/0102144.
- [7] V.P. Maslov. *Méthodes opératorielles*. Éditions MIR, Moscow, 1987.
- [8] V.P. Maslov and V.N. Kolokoltsov. *Idempotent analysis and its applications*. Kluwer Acad. Publ., Dordrecht et al., 1997.
- [9] G.L. Litvinov, V.P. Maslov, and G.B. Shpiz. Idempotent functional analysis. An algebraical approach. *Math. Notes*, 69(5):696–729, 2001. E-print arXiv:math.FA/0009128.
- [10] G. Litvinov and G. Shpiz. Dequantization of mathematics and group representations. *Vestnik Tamb. Univ.*, 10(4):390–411, 2005. in Russian.
- [11] G.L. Litvinov and A.N. Sobolevskii. Exact interval solutions to the discrete Bellman equation and polynomial complexity of problems in interval idempotent linear algebra. *Doklady Mathematics*, 62(2):199–201, 2000.
- [12] G. Litvinov and A. Sobolevskii. Idempotent interval analysis and optimization problems. *Reliable Computing*, 7(5):353–377, 2001. E-print arXiv:math.NA/0101152.

- [13] G. Litvinov and V. Maslov, editors. *Idempotent Mathematics and Mathematical Physics*, volume 377 of *Contemporary Mathematics*. American Mathematical Society, Providence, 2005.
- [14] F.L. Baccelli, G. Cohen, G.J. Olsder, and J.P. Quadrat. *Synchronization and Linearity*. Wiley, Chichester, New York, 1992.
- [15] R.A. Cuninghame-Green. *Minimax Algebra*, volume 166 of *Lecture Notes in Economics and Mathematical Systems*. Springer, Berlin, 1979.
- [16] P. Loreti and M. Pedicini. An object oriented approach to idempotent analysis: Integral equations as optimal control problems. In [13], pages 187–208.
- [17] G.H. Golub and C. Van Loan. *Matrix computations*. The John Hopkins University Press, Baltimore and London, 1989.
- [18] J. Golan. *Semirings and their applications*. Kluwer, Dordrecht, 2000.
- [19] G. Birkhoff. *Lattice theory*. Providence, Rhode Island, 1967.