

# Solving linear systems by methods based on a probabilistic interpretation

**Citation for published version (APA):**

van Nunen, J. A. E. E., & Wessels, J. (1979). *Solving linear systems by methods based on a probabilistic interpretation*. (Memorandum COSOR; Vol. 7904). Technische Hogeschool Eindhoven.

**Document status and date:**

Published: 01/01/1979

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Department of Mathematics

PROBABILITY THEORY, STATISTICS, OPERATIONS RESEARCH, AND SYSTEMS THEORY GROUP

Memorandum COSOR 79-04

Solving linear systems by methods  
based on a probabilistic interpretation

by

Jo van Nunen\* and Jaap Wessels

\* Graduate School of Management, Delft

Eindhoven, September 1978

The Netherlands

# SOLVING LINEAR SYSTEMS BY METHODS BASED ON A PROBABILISTIC INTERPRETATION

BY

Jo van Nunen<sup>†</sup> and Jaap Wessels\*

## Abstract

In this paper it is demonstrated how the probabilistic concept of a stopping time in a random process may be used to generate an iterative method for solving a system of linear equations. Actually all known iterative approximation methods for solving linear equations are generated by various choices of a stopping time e.g. the point and block Jacobi methods, the point and block Gauss-Seidel Methods and overrelaxation methods are covered.

The probabilistic approach offers -in a natural way- the possibility of adapting the solution technique to the special structure of the problem. Moreover, posterior bounds for the solution are constructed, which lead to faster convergence of the approximations than with usual prior bounds.

## 1. Introduction

In this paper we are concerned with iterative approximation methods for the square system of linear equations

$$(1.1) \quad A x = b, \quad x \in \mathbb{R}^N$$

Rewriting the system as

$$(1.2) \quad x = b + Px$$

with  $P = I - A$ , where  $I$  is the identity matrix, allows for a probabilistic interpretation of the system and therefore allows for the construction of

probabilistically based solution techniques. These probabilistic interpretations are feasible if the  $(i,j)$ -th entry  $p(i,j)$  of  $P$  is larger than or equal to zero and

$$(1.3) \quad \sum_{j=1}^N p(i,j) \leq 1 \quad \text{for all } i = 1,2,\dots,N.$$

Hence we will develop our method of generating solution techniques for that case first and indicate generalizations afterwards.

The system (1.1) or (1.2) is uniquely solvable if the spectral radius of  $P$  is smaller than one.

The solution of (1.2)

$$(1.4) \quad x^* = A^{-1} b = (I - P)^{-1} b = \sum_{t=0}^{\infty} P^t b$$

may be interpreted as the total expected reward belonging to a stochastic process that is considered at discrete points in time  $t = 0,1,2,\dots$ : as being in one of the states  $1,2,\dots,N$  and where we receive  $b(i)$  every time the system is observed to be in state  $i$ . The state  $i$  corresponds to the  $i$ -th equation in the linear system and to the  $i$ -th coordinate of  $x$ .

Using the concept of stopping time for this process yields the possibility to compute the expected rewards for a particular set of paths (sequences of states), in the random process.

The choice of the stopping time will determine the particular set.

Iterating the computation of expected rewards over that set finally yields the solution of the system (1.2). This solution will be reached independently of the chosen stopping time. So each stopping time will generate an iterative procedure for solving a system of linear equations. More explicitly each stopping time  $T$  will induce an operator  $L_T$  that is contractive on  $\mathbb{R}^N$  and has the solution of (1.2) as unique fixed point.

Actually all known iterative approximation methods for solving systems of linear equations are generated by the various stopping times. E.g. the point and block Jacobi methods, the point and block Gauss-Seidel methods

and overrelaxation methods are covered.

The probabilistic approach, however, offers in a natural way the possibility to adapt the solution technique to the special structure of the problem. In Markov decisions theory it is common to incorporate in the approximation technique the concept of upper and lowerbounds (MacQueen) [2] ) for the solution of (1.2).

It will be shown that applying this concept will lead to faster convergency of the approximation algorithms. The convergence will not be determined by the norm of the operator  $L_T$  but by the second largest eigenvalue. In section 2 the probabilistic interpretations of

$$x = b + P x \quad x \in \mathbb{R}^N$$

will be presented. The concept of stopping time will be introduced in an informal way. In section 3 stopping times will be used to generate iterative approximation procedures, and examples will be described.

In order to be able to give the central proof about the operators  $L_T$  we will describe the concept of a stopping time more formally in section 3. Examples of operators generated by stopping times will be given.

Section 4 will be devoted to upper and lower bounds for the solution of (1.2). These bounds will be based on the n-th stage of the approximation process and hence are posterior bounds. It appears that these posterior bounds are usually much better than the common prior bounds.

## 2. Probabilistic concepts

As mentioned, we will make the following simplifying assumption for the introduction of our main notions

$$A x = b$$

may be written as

$$x = b + P x$$

with  $p(i,j) > 0$ ,  $\sum_{j=1}^N p(i,j) \leq 1$  and  $P$  has a spectral radius  $\rho < 1$ .

Under these assumptions, the square linear system has the unique solutions  $\sum_{t=0}^{\infty} P^t b = (I - P)^{-1} b$  and the problem is to find the sum of this series  $t = 0$

or an approximation to it.

An interpretation for this sum is apparent if we interpret the square  $(N \times N)$  matrix  $P$  as the matrix of transition probabilities of a Markov chain. Namely, suppose a system can be in one of the states  $1, 2, \dots, N$ ; at time epoch  $t = 0, 1, 2, \dots$ . If it is in state  $i$  at time  $t$  it either jumps to a state  $j \neq i$  at time  $t + 1$  or it remains in the same state  $i$ . If the system is in state  $i$  at time  $t$  the probability of a transition to state  $j$  at time  $t + 1$  is equal to  $p(i, j)$ . The probability of remaining in state  $i$  is  $p(i, i)$ . The probability  $1 - \sum_{j=1}^N p(i, j)$  may be viewed upon as the probability for blowing the system up at time  $t$ . This implies e.g. that the probability of a path  $i, j, k, l$  at times  $0, 1, 2, 3$  given that the system starts in state  $i$  is equal to  $p(i, j)p(j, k)p(k, l)$ .

Hence the probability of finding the system in state  $j$  at time  $t$  given the system starts in state  $i$  at time  $t = 0$  is equal to

$$\sum_{i_1, i_2, \dots, i_{t-1}} p(i, i_1)p(i_1, i_2) \dots p(i_{t-2}, i_{t-1}) \cdot P(i_{t-1}, j)$$

This form however is equal to the  $(i, j)$ -th component of the  $(N \times N)$  matrix  $P^t$ . If we interpret  $b(i)$  as the reward received at each visit of the system to state  $i$ , then  $(P^t b)(i)$  is the expected reward at time  $t$  when the system starts in state  $i$ . Moreover  $\sum_{t=0}^{\infty} P^t b$  is the vector of total expected rewards for the different starting states until the system is blown up. This interpretation for the solution of (1.2) gives a good key to several approximation ideas. In the most simple procedure we start with

$$\begin{aligned} x_0 &= 0 & (x_0 \in \mathbb{R}^N) \\ x_1 &= b + P x_0 = b \end{aligned}$$

so,  $x_1(i)$  is the immediate reward we get if we start in state  $i$ .

Then we define

$$x_2 = b + P x_1 = b + P b$$

or componentwise

$$x_2(i) = b(i) + \sum_{j=1}^N p(i,j) b(j)$$

$x_2(i)$  is the expected reward in the first two periods given that the system starts in state  $i$ .

In general this simple procedure proceeds as follows

$$x_0 := 0 \quad x_n := b + P x_{n-1} \quad \text{for } n = 1, 2, \dots$$

We see that  $x_n$  is just the vector of expected rewards before time  $n$  for the different starting states. Using the same procedure with  $x_0 = a$  the vector  $x_n$  gives the expected rewards if the system is stopped (if not blown up earlier) at time  $n$  and a reward  $a(j)$  is earned if this occurs in state  $j$ . The vector  $a$  may be called a vector of terminal rewards.

Other procedures may be generated using so called stopping times. A stopping instruction for the process is an instruction which may use the path of the process until time  $t$  in order to decide whether the process will be stopped at  $t$  or not.

The random time at which the process is stopped according to a stopping instruction is called the stopping time.

Simple examples of stopping instructions are

- Example 2.1     stop at  $t = 1$
- Example 2.2     stop at  $t = 2$
- Example 2.3     stop as soon as the system reaches a state that is different from the starting state.
- Example 2.4     stop as soon as the system reaches a state with a higher number than the previous state
- Example 2.5     stop as soon as the system reaches a certain subset of the state space.

We will allow stopping instructions which require a lottery to make the stopping decision. As a very simple example we have

Example 2.6 Stop certainly as soon as the system reaches a state different from the starting state, but each time the system remains in the starting state the process is stopped with probability 1/2.

More sophisticated examples of randomized stopping instructions will be given later on. For a formal treatment of the concept of (randomized) stopping instructions, in a more general setting, see van Nunen [4].

3. The generation of iterative approximation schemes by stopping times.

Let us suppose for the moment that a stopping instruction is given. Denote the random stopping time by  $T$ . Then we can compute the total rewards until the process is stopped by

$$\sum_{t=0}^{T-1} b(S_t)$$

Where  $S_t$  denotes the random state of the system at time  $t$ . If a terminal reward  $a(j)$  is earned in case the process is stopped in  $j$ , the total rewards become

$$\sum_{t=0}^{T-1} b(S_t) + a(S_T).$$

(If  $T = \infty$  the final term is regarded as zero) The expected rewards given the starting state  $i$  are denoted by

$$\mathbf{E}_i \left[ \sum_{t=0}^{T-1} b(S_t) + a(S_T) \right]$$

the index  $i$  is deleted if the vector of such expectations is meant.



Now we consider the following procedure

$$\begin{aligned}
 & x_0 = a \\
 (3.1) \quad & x_n = \mathbb{E} \left[ \sum_{t=0}^{T-1} b(S_t) + x_{n-1}(S_T) \right] \quad n = 1, 2, \dots
 \end{aligned}$$

Note that for  $T \equiv 1$  (example 2.1) this procedure is exactly the same as the procedure introduced earlier.

The question remains for which types of stopping instructions do we have

$$(3.2) \quad \lim_{n \rightarrow \infty} x_n = [I - P]^{-1} b \quad ?$$

Intuitively this question is simple to answer:  $x_n(i)$  are the expected total rewards over  $n$  iterations of the stopping time plus an expected terminal reward at the end of this period given that we started in state  $i$ . If this period tends to infinity for  $n \rightarrow \infty$  then  $x_n(i)$  will tend to the total expected reward over the infinite time horizon.

As usual we may define the iterative procedure (3.1) by an operator  $L_T$

$$(3.3) \quad L_T x = \mathbb{E} \left[ \sum_{t=0}^{T-1} b(S_t) + x(S_T) \right].$$

The question in (3.2) now becomes: does it hold that

$$(3.4) \quad \lim_{n \rightarrow \infty} L_T^n a = (I - P)^{-1} b \quad \text{for all } a \in \mathbb{R}^N$$

with  $L_T^{n+1} = L_T(L_T^n)$

This question can be split in two parts

- (i) is  $(I - P)^{-1} b$  a fixed point of  $L_T$  ?
- (ii) is  $L_T$  a contracting operator ?

Then (3.4) holds if both questions can be answered in a positive way.

The answer to the first question happens to be yes because

$$(3.5) \quad L_T \left( \sum_{n=0}^{\infty} P^n b \right) = \mathbb{E} \left[ \sum_{t=0}^{T-1} b(S_t) + \left( \sum_{n=0}^{\infty} P^n b \right) (S_T) \right]$$

Where the infinite series in the right handside can be interpreted as the expected reward from the stopping time on if the instruction to stop in state  $S_T$  at time  $T$  is ignored. So the whole expression is equal to the total expected reward if there was no stopping instruction. This total expected reward was just the solution of the system of linear equations (1.2). For a formal proof see van Nunen [4] or Wessels [10].

The answer to the second question cannot be an unconditional one since e.g. the case  $T \equiv 0$  gives trivially  $L_T x = x$  for any  $x$ .

Even the most straight forward case  $T \equiv 1$  requires some subtlety, since  $P$  need not be contracting in maximum norm and therefore  $L_T x = b + Px$  need not be contracting.

However  $P$  is certainly  $N$ -stage contracting (for  $x$ ) since its spectral radius is less than one and therefore  $L_T$  with  $T \equiv 1$  is  $N$ -stage contracting and possesses a unique fixed point  $(I-P)^{-1}b$ .

For other situations we have for any stopping time  $T$ .

$$(3.6) \quad (L_T x)(i) = \mathbb{E}_i \left[ \sum_{t=0}^{T-1} b(S_t) \right] + \sum_j P_{ij}(S_T = j, T < \infty) x(j)$$

where  $P_{ij}(S_T = j, T < \infty)$  is the probability that the state at the stopping time  $T$  equals  $j$  and  $T < \infty$  for given starting  $i$ .

This implies that  $L_T$  is  $N$ -stage contracting analogous to the case  $T \equiv 1$ , if the matrix with  $(i,j)$ -th entry  $P_{ij}(S_T = j, T < \infty)$  is  $N$ -stage contracting. We will prove that this holds if and only if the stopping instruction is such that for any starting state the probability of immediate stopping is strictly less than one. Such a stopping instruction, that does not allow for immediate stopping at  $t = 0$  with probability one will be called a nonzero stopping instruction and the corresponding stopping time will be referred to as a nonzero stopping time.

Theorem 1 For any, possibly randomized, nonzero stopping time the operator  $L_T$  is  $N$ -stage contracting and has the unique fixed point

$$x = (I - P)^{-1} b = \sum_{t=0}^{\infty} P^t b .$$

In order to be able to prove the above theorem we have to prove that the operator  $L_T$  is N-stage contracting if T is a nonzero stopping time. In order to treat this problem properly we have to introduce the stopping instructions in a more formal way.

Definition A (randomized) stopping instruction is a function  $\delta$  from the set of all possible paths given by  $G_\infty = \bigcup_{k=1}^{\infty} S^k$  (with  $S = \{1, 2, \dots, N\}$  and  $S^k = S \times S \times \dots \times S$ ) into  $[0, 1]$ .

For any sequence of states  $i_0, i_1, i_2, \dots, i_k$  the number of  $\delta(i_0, \dots, i_k)$  will give the probability that the process is stopped in state  $i_k$  at time k given that the path until time k was  $i_0, \dots, i_k$  and that the process has not been stopped earlier.

Remarks

3.1 Note that if for all  $\alpha \in G_\infty$  we have  $\delta(\alpha) \in \{0, 1\}$ , then the stopping instruction is nonrandomized.

3.2 The stopping instruction  $\delta$  is nonzero if and only if  $\delta(i) \neq 1$  for all  $i \in S = \{1, 2, \dots, N\}$ .

Examples

3.1 The stopping instruction defined in example 2.1 is nonzero and equals  $\delta(i) = 0$  for all  $i \in S$  and  $\delta(\alpha) = 1$  for all  $\alpha \in G_\infty \setminus S$ .

3.2. The stopping instruction defined in ex. 2.2 equals  $\delta(\alpha) = 0$  for  $\alpha \in S \cup S^2$  and  $\delta(\alpha) = 1$  for  $\alpha \in G_\infty \setminus (S \cup S^2)$ .

3.3 The stopping instruction described in ex. 2.3 has  $\delta(\alpha) = 0$  if and only if all components of  $\alpha \in G_\infty$  are equal e.g.  $\alpha = (3, 3, 3, \dots, 3)$  and  $\delta(\alpha) = 1$  else.

3.4 The stopping instruction in ex. 2.4 has  $\delta(i_0, i_1, \dots, i_k) = 0$  iff  $i_0 \geq i_1 \geq \dots \geq i_k$  in all the other situations  $\delta(\alpha) = 1$ .

3.5 Example 2.5 corresponds to  $\delta(\alpha) = 0$  iff  $\alpha = (i_0, i_1, \dots, i_a)$  and  $i_0, i_1, \dots, i_a \in E$  where  $E$  is the special stopping set and  $\delta(\alpha) = 1$  else.

3.6 Example 2.6 corresponds to  $\delta(\alpha) = \frac{1}{2}$  iff  $\alpha = (i_0, i_1, \dots, i_k)$  and  $i_0 = i_1 = \dots = i_k$ ;  $\delta(\alpha) = 1$  else.

Note that 3.1 - 3.5 give nonrandomized stoppings instructions while 3.6 is randomized. All examples give nonzero stopping times.

Before proving that  $L_T$  is  $N$ -stage contracting, let us first introduce some notations. The operator  $L_T$  defined in (3.6) can be rewritten as

$$(3.7) \quad L_T x = b_T + P_T x$$

with

$$b_T = \mathbb{E} \sum_{t=0}^{T-1} b(S_t)$$

and  $P_T$  is the  $(N \times N)$ -matrix with  $(i, j)$ -th entry equal to

$$\mathbb{P}_i (S_T = j, T < \infty)$$

This in turn is equal to

$$\sum_{n=0}^{\infty} \mathbb{P}_i (S_n = j, T = n) .$$

We will prove our point under the more restrictive assumptions that  $\|P\| < 1$  where  $\|P\| = \max_i \sum_j |p(i, j)|$ . The extension to the case where  $P$  is  $N$ -stage contracting is then straight forward.

From (3.7) it appears that it suffices to show the contractingness of  $P_T$ .

Lemma 1 Let  $\delta$  be a nonzero stopping instruction and suppose  $\|P\| < 1$ , then

$$\rho_T := \|P_T\| \leq \max_i \delta(i) + (1 - \max_i \delta(i)) \|P\| < 1$$

Proof First we note that  $\|P_T\|$  is finite, since  $\|P_T\| \leq \sum_{n=0}^{\infty} \|P^n\| < \infty$ ,  
(  $\|P\| < 1$  )

The proof will be given by induction. For the stopping instructions  $\delta$  we define the instruction  $\delta_M$  by

$$\delta_M(\alpha) := \begin{cases} 1 & \text{if } \alpha \in \bigcup_{k=M+1}^{\infty} S^k \\ \delta(\alpha) & \text{else} \end{cases}$$

Then for every  $\varepsilon > 0$  there exists a  $K \in \mathbb{N}$  such that for all  $n > K$

$$| \| P_{T_M} \| - \| P_T \| | < \varepsilon$$

with  $P_{T_M}$  and  $P_T$  the matrices corresponding to the stopping instructions  $\delta_M$  and  $\delta$  respectively.

The above inequality is true since

$$| \| P_{T_M} \| - \| P_T \| | \leq \sum_{n=M}^{\infty} \| P^n \|$$

So it suffices to prove the lemma for nonzero "finite" stopping instructions of the type  $\delta_M$ .

Let  $\Delta_n$  be the set of all stopping instructions for which the process is stopped with certainty before or at time  $n$  to

$$\Delta_m = \{ \delta \mid \delta(\alpha) = 1 \text{ for all } \alpha \in \bigcup_{k=n+1}^{\infty} S^k \}$$

Consequently  $\Delta_0$  only contains the stopping instructions that require immediate stopping

i.e.  $\delta(i) = 1$  for  $i = 1, \dots, N$ , which implies  $\| P_\delta \| = 1$ .

Suppose  $\delta \in \Delta_1$  is a nonzero stopping instruction then

$$\begin{aligned} \sum_{j \in S} P_T(i, j) &= \sum_{j \in S} [P_i(s_0=j, T=0) + P_i(s_1=j, T=1)] \\ &= \delta(i) + (1 - \delta(i)) \sum_{j \in S} p(i, j) \\ &\leq \delta(i) + (1 - \delta(i)) \| P \| \end{aligned}$$

Since  $\delta \in \Delta_1$  is supposed to be nonzero  $\delta(i) \leq 1 - \varepsilon$  for some  $\varepsilon > 0$  and

$\|P\| \leq 1$ , we have  $\|P_T\| < 1$  for  $T$  corresponding to  $\delta \in \Delta_1$ .

Now we state the induction hypothesis. Let  $\delta \in \Delta_n$  be any stopping instruction and  $T$  the stopping time corresponding to  $\delta$ .

Suppose

$$\|P_T\| < 1 \quad \text{and} \quad \|P_T\| < 1 \quad \text{if } \delta \text{ is nonzero.}$$

For  $\delta' \in \Delta_{n+1}$  we define  $\delta_i(\alpha) = \delta(i, \alpha)$  for  $i \in S$  and  $\alpha \in \prod_{k=1}^{\infty} S^k$ . Then clearly  $\delta_i \in \Delta_n$ .

Now for any  $i \in S$  we have

$$\begin{aligned} \sum_{j \in S} P_T(i, j) &= \sum_{j \in S} \sum_{m=0}^{n+1} P_i(S_m = j, T = m) \\ &= \sum_{j \in S} [ P_1(S_0 = j, T=0) + \sum_{m=1}^{n+1} P_i(S_m = j, T=m) ] \\ &= \delta(i) + (1-\delta(i)) \sum_{j \in S} \sum_{m=1}^{n+1} \sum_{k \in S} p(i, k) P_i(S_m = j, T=m | S_1 = k) \\ &= \delta(i) + (1-\delta(i)) \sum_{k \in S} p(i, k) \sum_{j \in S} \sum_{k=0}^n P_k(S_m = j, T_i = m) \\ &\leq \delta(i) + (1-\delta(i)) \sum_{k \in S} p(i, k) \\ &\leq \delta(i) + (1-\delta(i)) \|P\| \end{aligned}$$

Where the last but one inequality follows from the induction hypothesis.

So for nonzero  $\delta \in \Delta_n$

$$\|P_T\| = \rho_T < 1.$$

Corollary 1 For any, possibly randomized, nonzero stopping time the sequence  $x_n$  defined by

$$x_n = L_T x_{n-1} = (L_T)^n x_0$$

converges to the solution of the system  $(I-P)x = b$ .

From the foregoing it follows that every nonzero stopping time  $T$  defines an operator  $L_T$  which in turn induces a successive approximation algorithm for solving the system of linear equations (1.1). For solving a particular problem, one has to exploit the specific structure in order to find the best stopping time within the class of all stopping times.

Examples

3.1.a Let the stopping instruction be as in ex.3.1 and let  $x_0 \in \mathbb{R}^N$  then

$$x_{n+1}^{(i)} := (L_T x_n)(i) = b(i) + \sum_{j=1}^N p(i,j)x_n(j)$$

In matrix notation

$$x_n = b + P x_{n-1}$$

This procedure is sometimes called the pre-Jacobi method.

3.3.a Let the stopping instruction be as in example 3.3

Then for  $x_0 \in \mathbb{R}^N$  the sequence  $x_n$  is defined by

$$(3.8) \quad x_{n+1}^{(i)} = (L_T x_n)(i) = \frac{1}{1-p(i,i)} \left[ b(i) + \sum_{j \neq i} p(i,j)x_n(j) \right]$$

This follows since the stopping instruction guarantees the reception of  $b(i)$  everytime we remain in state  $i$  for one more period. So we earn  $b(i)$  in first instance then we earn again  $b(i)$  in the next period with probability  $p(i,i)$ . After 2 steps we are still in state  $i$  with probability  $(p(i,i))^2$ . So the expected contribution is  $(p(i,i))^2 b(i)$  for step 2. In step 3 the contribution will be  $(p(i,i))^3 b(i)$ . Continuing in this way yields the geometric series that sums up to

$$\frac{1}{1-p(i,i)} b(i)$$

The second term in  $x_{n+1}$  follows from a similar reasoning. The process is stopped after one period in state  $j$  with probability  $p(i,j)$ . ( $j \neq i$ ) and the final reward  $x_n(j)$  is earned. The process is stopped after 2 periods in state  $j \neq i$  iff it was still in state  $i$  after one period. This happens

with probability  $p(i,i)p(i,j)$ . The process is stopped in  $j \neq i$  after 3 transitions with probability  $(p(i,i))^2 p(i,j)$ . Continuing this concept leads to the second term in the righthand side of (3.8).

We define  $D$  as the diagonal matrix with diagonal entries  $(1-p(i,i))$  and the  $(N \times N)$ -matrices  $F$  and  $E$  as the strictly upper and lower triangular parts of  $P$  respectively (and zero entries elsewhere).

Then (3.8) can be reformulated as

$$x_{n+1} = D^{-1}(E + F)x_n + D^{-1} b$$

This is the point Jacobi or point total step method see e.g. Varga [9].

3.4.a Let the stopping instructions be as in example 3.4 then

$$(3.9) \quad x_{n+1}(i) := \frac{1}{1-p(i,i)} b(i) + \frac{1}{1-p(i,i)} \sum_{j < i} p(i,j) x_{n+1}(j) \\ + \frac{1}{1-p(i,i)} \cdot \sum_{j > i} p(i,j) x_n(j)$$

This follows easily by noting that the reward earned after a transition to state  $j < i$  equals exactly  $x_{n+1}(j)$ .

By using matrices  $D$ ,  $E$  and  $F$  as defined in the previous example, (3.9) can be given by

$$x_{n+1} = (D-E)^{-1} F x_n + (D-E)^{-1} b$$

This iterative procedure is known as the point Gauss-Seidel or point single-step method. See e.g. Varga [9].

Ex. 3.7 Let the stopping instruction be such that we stop after  $i_0, i_1, \dots, i_n, i$  in state  $i$  with probability  $\delta(i)$  if

$$i_0 \geq i_1 \geq \dots \geq i_n \geq i,$$

in other situations we stop immediately ( $\delta(\alpha) = 1$ ).



Then in similar way as in the previous example we have

$$\begin{aligned}
 (3.10) \quad x_{n+1}(i) &= \frac{1-\delta(i)}{1-(1-\delta(i))p(i,i)} b(i) + \frac{\delta(i)}{1-(1-\delta(i))p(i,i)} x_n(i) \\
 &+ \frac{1-\delta(i)}{1-(1-\delta(i))p(i,i)} \sum_{j<i} p(i,j) x_{n+1}(j) + \\
 &+ \frac{1-\delta(i)}{1-(1-\delta(i))p(i,i)} \sum_{j>i} p(i,j) x_n(j)
 \end{aligned}$$

Let us define  $\delta(i)$  such that  $\frac{(1-\delta(i))(1-p(i,i))}{1-(1-\delta(i))p(i,i)} = \text{constant} = \omega$

Then - again using the matrices D, E and F (3.10) becomes

$$(D - \omega E)x_{n+1} = \omega b + [(1-\omega)D + \omega F] x_n$$

which is known as the point successive overrelaxation method.

Ex. 3.8 Suppose the state space is partitioned in blocks.

$B_1 = \{1, 2, \dots, n\}$  ;  $B_2 = \{n_1+1, \dots, n_2\}$  , ...,  $B_k = \{n_{k-1}+1, \dots, N\}$ .

$$P := \begin{bmatrix} P_{1,1} & P_{1,2} & \dots & P_{1,k} \\ P_{2,1} & P_{2,2} & \dots & P_{2,k} \\ \vdots & & & \vdots \\ P_{k,1} & \dots & \dots & P_{k,k} \end{bmatrix}$$

Where  $P_{nm}$  contains the entries  $p(i,j)$  with  $i \in B_n$  and  $j \in B_m$ . Let the  $(N \times N)$ -matrices D, E, F be defined by

$$\begin{aligned}
 D &:= I - \begin{bmatrix} P_{11} & & 0 \\ & P_{22} & \\ 0 & & P_{kk} \end{bmatrix} \\
 E &:= \begin{bmatrix} 0 & & & 0 \\ P_{2,1} & & & \\ & \ddots & & \\ P_{k,1} & & & P_{k,k-1} & 0 \end{bmatrix} \\
 F &:= \begin{bmatrix} 0 & P_{1,2} & & & P_{1,k} \\ & 0 & & & \\ & & \ddots & & \\ & & & 0 & P_{k-1,k} \\ 0 & & & & 0 \end{bmatrix}
 \end{aligned}$$

Now let  $\delta$  be the stopping instruction with  $\delta(i) = 0$  for all  $i \in S$ ,

$$\delta(i_0, i_1, \dots, i_j) = 0 \text{ if } i \text{ and } j \text{ are in the same block else } \delta(\alpha) = 1$$

This means that the system is not stopped as long as the states that are "visited" belong to the same block. Now it is straight forwardly verified that

$$x_{n+1} = L_T x_n \quad \text{can be reformulated as}$$

$$x_{n+1} = D^{-1}(E + F)x_n + D^{-1}r$$

This iterative scheme is known as the block Jacobi iterative method (see Varga [9]).

From the foregoing examples it will be clear that the so-called block successive overrelaxation iterative method can be formed by combining the examples 3.7 and 3.8.

Of course several other options are available. The previous examples are only given to illustrate how the concept of stopping time can be used to generate iterative methods for solving systems of linear equations. For specially structured matrices one can choose the (randomized) stopping instruction in such a way that the relevant paths are included by the mapping  $L_T$ .

Remarks

- a. To apply the idea of stopping instructions it is not necessary that all non-diagonal elements of the matrix A in (1,1) are non-positive and that  $\sum_{j=1}^N P(i,j) \leq 1$ . In fact we only used these restrictions in order to obtain a probabilistic interpretation of the system of equations. For the formal proofs it suffices that A is nonsingular and that  $P^n \rightarrow 0$  for  $n \rightarrow \infty$ .
- b. In fact, the deduction of successive approximation methods by using the concept of stopping time leads to a regular splitting of the (NxN)-matrix A, (See again Varga [9]).  
The real matrices Q, R with  $A := Q - R$  form a regular splitting of A if Q is nonsingular with  $Q^{-1} > 0$  and  $R \geq 0$ . If, in addition,  $A^{-1} \geq 0$ , then the sequence  $x_{n+1} = Q^{-1}R x_n + Q^{-1}b$  converges to the solution of  $Ax = b$ .  
Where  $Q^{-1}R$  corresponds to  $P_T$  and  $Q^{-1}b = E \sum_{t=0}^{T-1} b(S_t)$ .  
This relation -although not formulated in stopping times- can also be found in Porteus [7].
- c. The theory of stopping times for Markov chains gives us the following result which is probabilistically trivial.

Lemma 3.2 (see van Nunen [4] p. 31)

Let  $\delta_1$  and  $\delta_2$  be stopping instructions such that  $\delta_1(\alpha) \geq \delta_2(\alpha)$  for all  $\alpha \in \bigcup_{k=1}^{\infty} S^k$  then

$$\rho(P_{T_1}) \geq \rho(P_{T_2})$$

where  $\rho$  indicates the spectral radius.

Essentially the same lemma can be found in text books of numerical mathematics.

Lemma 3.2a (see Varga [9] pp. 81)

If  $A := Q_1 - R_1 = Q_2 - R_2$  are two regular splittings of  $A$  where  $A^{-1} \geq 0$  and  $R_2 \geq R_1 \geq 0$  then

$$0 \leq \rho(Q_2^{-1} R_2) \leq \rho(Q_1^{-1} R_1)$$

As a consequence of these lemmas it follows immediately that e.g. the spectral radius for the Gauss-Seidel method is not larger than the spectral radius for the Jacobi method. However it needs not be true that the bounds that can be constructed (see section 4) are better too (van Nunen [4]).

#### 4. Upper and lower-bounds for the solution of $Ax = b$ .

In the previous section we showed that each stopping instruction generates a sequence  $\{x_n\}_{n=0}^{\infty}$  that converges to the solution of (1.1).

By examples we illustrated that these sequences cover those arising from the methods which are well known from numerical mathematics.

We did not give much attention to the convergence rate of the methods.

In this section the emphasis will be on the rate of convergence. We will show how a faster convergence rate can be achieved by using the concept of posterior upper and lowerbounds for the solution  $x^*$  of (1.1).

For a sequence  $\{x_n\}$  we define  $\alpha_n$  and  $\beta_n$  for  $n = 1, 2, \dots$  by

$$\alpha_n = \min_i [x_n(i) - x_{n-1}(i)],$$

$$\beta_n = \max_i [x_n(i) - x_{n-1}(i)].$$

Moreover, we define for a matrix  $B$  two types of numbers

$$\|B\| = \max_i \sum_j B(i, j)$$

$$|B| = \min_i \sum_j B(i, j)$$

These numbers will be used in defining the upper and lowerbounds for the solution  $x^*$ . Note that  $\|B\|$  is the norm of  $B$  in maximum norm sense if  $B \geq 0$ .

Lemma 4.1

Let  $\delta$  be any nonzero stopping instruction and  $x_0 \in \mathbb{R}^N$ . Let the sequence  $\{x_n\}$  be defined by means of the stopping time based on the stopping instruction  $\delta$  as in (3.3)

$$x_n = L_T x_{n-1}.$$

Then

(i) The sequence  $\{u_n\}_{n=1}^{\infty}$  defined by

$$u_n = \begin{cases} x_n + \|P_T\| (1 - \|P_T\|)^{-1} \beta_n e, & \text{if } \beta_n \geq 0 \\ x_n + |P_T| (1 - |P_T|)^{-1} \beta_n e & \text{if } \beta_n < 0 \end{cases}$$

yields monotone non-increasing upperbounds for  $x^n$ , if  $e$  denotes the transpose of  $(1, 1, \dots, 1)$ .

Moreover:  $u_n \downarrow x^*$ .

(ii) The sequence  $\{b_n\}_{n=1}^{\infty}$  defined by

$$b_n = \begin{cases} x_n + |P_T| (1 - |P_T|)^{-1} \alpha_n e, & \text{if } \alpha_n \geq 0 \\ x_n + \|P_T\| (1 - \|P_T\|)^{-1} \alpha_n e, & \text{if } \alpha_n < 0 \end{cases}$$

$$(iii) \quad \|u_n - l_n\| \leq 2 \|P_T\|^{n+1} (1 - \|P_T\|)^{-1} \|x_1 - x_0\|$$

$$\|u_n - l_n\| \leq 2 \|P_T\| (1 - \|P_T\|)^{-1} \|x_n - x_{n-1}\|$$

*Proof* For a complete proof we refer to van Nunen [4]. Similar results have been found by Porteus [7]. However, the underlying ideas for the proof are rather simple i.e.

$$x_{n+1} - x_n = L_T x_n - L_T x_{n-1} = P_T (x_n - x_{n-1})$$

$$\leq \begin{cases} \|P_T\| \beta_n e, & \text{if } \beta_n \geq 0 \\ |P_T| \beta_n e, & \text{if } \beta_n < 0 \end{cases}$$

Using the analogous inequality for  $L_T^2, L_T^3$  etc. yields

$$x_{n+k} - x_n = (x_{n+k} - x_{n-k-1}) + (x_{n-k-1} - x_{n-k-2}) + \dots + (x_{n+1} - x_n)$$

$$\leq \begin{cases} (\|P_T^k\| + \|P_T^{k-1}\| + \dots + \|P_T\|) \beta_n e & \text{if } \beta_n \geq 0 \\ (|P_T^k| + |P_T^{k-1}| + \dots + |P_T|) \beta_n e & \text{if } \beta_n < 0 \end{cases}$$

Since  $\|\cdot\|$  is a norm and hence satisfies  $\|CD\| \leq \|C\| \cdot \|D\|$  for nonnegative  $C, D$  and since  $|\cdot|$  satisfies  $|CD| \geq |C| \cdot |D|$  for nonnegative  $C, D$ , we obtain for  $k \rightarrow \infty$

$$x^* \leq \begin{cases} x_n + \|P_T\| (1 - \|P_T\|)^{-1} \beta_n e & \text{if } \beta_n \geq 0 \\ x_n + |P_T| (1 - |P_T|)^{-1} \beta_n e & \text{if } \beta_n < 0 \end{cases}$$

The other statements can be found in a similar way.

From part (iii) of the foregoing lemma it follows that the convergence of the sequence

$$x_n' := \frac{1}{2} (u_n + l_n)$$

is at least at a rate  $\|P_T\|$ .

We will show that the use of bounds may lead to a nonnegligeable acceleration of the convergence.

Moreover, the asymptotic convergence rate may be improved.

Although this is not necessarily requested to derive the results in the sequel of this section we will -for reasons of simplicity- only consider the following simple situation.

Assumption 4.1

(i)  $\alpha_1 \geq 0$

(ii)  $P_T$  has the equal row sum property i.e.

$$\sum_j p(i,j) = \beta < 1 \text{ for all } i.$$

The first assumption implies that we can just use the upper parts of the definitions for  $u_n, l_n$  in lemma 4.1.(i) and 4.1(ii), since  $\alpha_1 \geq 0$  implies

$$\alpha_n \geq 0 \text{ for } n \geq 1 \text{ and hence } \beta_n > 0.$$

The condition  $\alpha_1 \geq 0$  can always be satisfied by choosing  $x_0$  in an adequate way.

Note that  $\alpha_n \geq 0$  implies that the convergence is monotone i.e.

$$x_{n-1} \leq x_n \leq x^*$$

Consequently

$$x_n \leq l_n \leq x^* \leq u_n$$

The second assumption can also be imposed without loss of generality.

Any problem with a  $P_T$  matrix that has unequal row sums can be transformed into an equivalent problem with a  $P'_T$  matrix which possess this equal row sum property. This transformation is due to Schweitzer [12]. It can be found in Porteus [7] and may also be deduced by using a stopping time (see van Nunen and Stidham [6]).

The corresponding stopping instruction  $\delta_{T \circ T'}$  for this stopping time is based on a combination of  $\delta_T$  and the instruction  $\delta'$  defined by  $\delta'(i) = 0$  for all

$$i \text{ and } \delta'(i,i) = 1 - \frac{1}{p_T(i,i)} (\sum_j p_T(i,j) - \gamma)(1 - \gamma) \text{ with}$$

$$\gamma = \max_i (\sum_{j \neq i} p_T(i,j) \cdot (1 - p_T(i,i))^{-1}).$$

Now, a straight forward verification shows that the matrix  $P_{T \circ T'}$  has all row sums equal to  $\gamma < 1$ , where  $T \circ T'$  corresponds to the stopping instruction that combines  $T$  and  $T'$ .

From lemma 3.2 it follows that  $\|P_{T \circ T'}\| \leq \|P_T\|$ .

In the situation that all row sums of  $P_T$  are equal we have of course  $\rho(P_T) = \|P_T\| = |P_T|$ .

Consequently we have for that situation

$$\begin{aligned} u_n - l_n &= \frac{\|P_T\|}{1 - \|P_T\|} (\beta_n - \alpha_n) e \\ &= \frac{\|P_T\|}{1 - \|P_T\|} \cdot \text{sp}(x_n - x_{n-1}) \end{aligned}$$

where  $\text{sp}(x) := \max_i x(i) - \min_i x(i)$

Lemma 4.2

Assume  $P_T$  has the equal row sum property, then

- (i) the convergence rate of the procedure based on  $P_T$  is bounded by the sub dominant eigenvalue  $\lambda$  of  $P_T$  since  $\text{sp}(P_T x) \leq |\lambda| \cdot \|P_T\| \text{sp}(x)$

- (ii) an upperbound for  $|\lambda_T|$  can be given by

$$v_T = \max_{i,j} \left[ \sum_{l=1}^1 \frac{1}{\|P_T\|} \max(P_T(i,l) - P_T(j,l), 0) \right] \leq 1$$

Proof

For the proof of part (i) we refer to Morton and Wecker [3].

The proof of part (ii) can be found e.g. in Seneta [8] or Hübner [1] or Morton and Wecker [3] and uses the following result



$$\begin{aligned}
 \text{sp}(P_T x) &= \max_i \left[ \sum_l [p_{T(i,l)} x(l)] \right] - \min_j \left[ \sum_l [p_{T(j,l)} x(l)] \right] \\
 &< \|P_T\| \cdot \left\{ \frac{1}{\|P_T\|} \max_{i,j} \left( \sum_l \max [p_{T(i,l)} - p_{T(j,l)}, 0] \right) \right\} \cdot \text{sp}(x) \\
 &= \|P_T\| \cdot \gamma_T \cdot \text{sp}(x).
 \end{aligned}$$

This implies that, although the norm and spectral radius of  $P_T$  are equal to  $\|P_T\|$ , (if  $P_T$  has equal row sums), the convergence rate of the sequence  $\frac{1}{2} (u_n + l_n)$  equals  $\|P_T\| \cdot \gamma_T$ .

To illustrate the strength of the concept of posterior upper and lower bounds we consider the following simple example.

$$P = .9 \begin{pmatrix} 0 & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} & 0 \end{pmatrix} \quad b = \begin{pmatrix} 10 \\ 8 \\ 0 \end{pmatrix}$$

Let  $x_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ , clearly  $\|P\| = .9$  while  $\gamma = .75$ .

In the next three tables we give for the pre-Jacobi iterates: the values  $x_n$ ;  $\frac{1}{2} (u_n + l_n)$  and  $\text{sp}(x_n - x_{n-1})$  for several values of  $n$  respectively.

TABEL 1  
pre-Jacobi iterates with  $x_0^t = (0,0,0)$

$i \backslash n$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_{10}$	$x_{15}$	$x_\infty = x^*$
1	10.00	12.40	18.34	21.84	25.76	39.34	47.39	59.00
2	8.00	12.50	17.02	21.31	24.86	38.55	46.59	58.20
3	0.00	7.65	11.23	15.62	19.30	32.90	40.95	52.56

Table 2 shows the values for the sequences  $\frac{1}{2}(u_n + l_n)$ .

TABLE 2

$$a_n = \frac{1}{2}(u_n + l_n)$$

n \	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>	a <sub>10</sub>
i=1	55.00	51.62	61.18	37.35	57.37	59.00
i=2	53.00	57.72	59.86	56.81	58.48	58.20
i=3	45.00	52.87	54.07	51.11	52.92	52.56

TABLE 3

$$d_n = sp(x_n - x_{n-1})$$

n \	d <sub>1</sub>	d <sub>2</sub>	d <sub>3</sub>	d <sub>4</sub>	d <sub>5</sub>	d <sub>10</sub>
i=1	10	2.40	5.94	3.50	3.92	2.19
i=2	8	4.50	4.52	4.28	3.55	2.19
i=3	0	1.65	3.58	4.39	3.68	2.18
sp	10	5.25	2.36	0.89	0.37	0.01

It is indicated by this very simple example that using the concept of bounds leads to very fast convergence. In the example the estimate was within 1% of the exact solution after 10 iterates.

## 5. Conclusions

We did not strive to be exhaustive in finding the best technique for solving a system of linear equations. We merely aimed to give more relations between the techniques used for solving Markov decision problems and those for linear systems.

Moreover, we showed how the concept of posterior lower and upperbounds, which is used in Markov decision theory can be exploited in solving linear systems.

There are of course a lot of improvements possible. For example column reductions or diagonal reductions (see e.g. Porteus [7]) can be used.

However it can be shown that these concepts do not improve the convergence rate once the concept of bounds is used.

The norm used in this paper was the maximum norm. The same concepts of stopping times and bounds can be used in the case where other norms are more suitable.

For an exposition on the use of weighted supremum norms for Markov decision processes we refer to Wessels [11] and van Nunen [4]

REFERENCES

- [1] Hübner, G." Improved procedures for eliminating suboptimal actions in Markov programming by the use of contraction properties." In: Transactions of the 7<sup>th</sup> Prague conference on stochastic processes, volume A, Academia prague (1977) pp.
- [2] MacQueen, J., "A modified dynamic programming method for Markovian decision problems" J. Math. Anal. appl. 14 (1966), pp.38-43.
- [3] Morton, T. and Wecker, W., "Discounting, ergodicity and convergence for Markov decision process." Man.Sci. 23 (1977) pp.890-900.
- [4] Nunen, J.A.E.E. van, "Contracting Markov decision processes". Mathematical centre tract no. 71. Math.Centre Amsterdam, 1976.
- [5] Nunen, J.A.E.E. van, "A set of successive approximation methods for discounted Markov decision problems." Zeitschrift für Oper.Res. 20, (1976), pp. 203-208.
- [6] Nunen, J.A.E.E. van and Stidham S. jr., "Action dependent stopping times and Markov decisions, processes with unbounded reward" Report: Program in O.R. N.C. state university. Raleigh N.C. (1978).
- [7] Porteus, E.L. "Bounds and transformations for discounted finite Markov decision chains." Oper.Res. 23 (1975), 761-784.
- [8] Seneta, E., "Nonnegative Matrices". Londen, Allyn and Bacon (1973).
- [9] Varga, R.S., "Matrix iterative analysis" Englewood Cliff, Prentice Hall 1962.
- [10] Wessels, J., "Stopping times and Markov programming". In: transactions of the 7<sup>th</sup> prague conference on stochastic programming etc. Volume A, Academia prague 1977. pp.575-585.

- [11] Wessels, J., "Markov programming by successive approximate with aspect to weighted supremum norm." J.Math.Anal.Appl. 58 (1977) pp.326-335.
- [12] Schweitzer, P. "Iterative solution of the functional equations of undiscounted Markov renewal programming." J.Math. Anal. appl. (1971) pp.495-501.