Journal of Computational and Applied Mathematics 227 (2009) 27–50

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Projected equation methods for approximate solution of large linear systems

Dimitri P. Bertsekas [a,*], Huizhen Yu [b]

[a] Department of Electr. Engineering and Comp. Science, M.I.T., Cambridge, MA, 02139, United States

[b] Helsinki Institute for Information Technology, University of Helsinki, Finland

ARTICLE INFO

Article history:
Received 26 February 2008
Received in revised form 29 May 2008

Keywords:
Linear equations
Projected equations
Dynamic programming
Temporal differences
Simulation
Value iteration
Jacobi method

ABSTRACT

We consider linear systems of equations and solution approximations derived by projection on a low-dimensional subspace. We propose stochastic iterative algorithms, based on simulation, which converge to the approximate solution and are suitable for very large-dimensional problems. The algorithms are extensions of recent approximate dynamic programming methods, known as temporal difference methods, which solve a projected form of Bellman's equation by using simulation-based approximations to this equation, or by using a projected value iteration method.

1. Introduction

In this paper we focus on systems of linear equations of the form

$$x = Ax + b, \tag{1.1}$$

where A is an $n \times n$ matrix and b is a column vector in the n-dimensional space $\Re^n$. We propose methods to compute an approximate solution within a subspace spanned by a relatively small number of basis functions.

Our motivation comes from recent advances in the field of dynamic programming (DP), where large systems of equations of the form (1.1) appear in the context of evaluation of the cost of a stationary policy in a Markovian decision problem. In this DP context, we are given an n-state Markov chain with transition probability matrix P, which evolves for an infinite number of discrete time periods, and a cost vector $g \in \Re^n$, whose components g_i represent the costs of being at the corresponding states $i = 1, \ldots, n$, for a single time period. The problem is to evaluate the total cost vector

$$x^* = \sum_{t=0}^{\infty} \alpha^t P^t g,$$

where $\alpha \in (0, 1)$ is a discount factor, and the components x_i represent the total expected α-discounted cost over an infinite number of time periods, starting from the corresponding states $i = 1, \ldots, n$. It is well known that x is the unique solution of the equation

$$x = \alpha Px + g,$$

* Corresponding author.
E-mail address: dimitrib@mit.edu (D.P. Bertsekas).

0377-0427/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2008.07.037

and furthermore, x can also be computed iteratively by the Jacobi method $x_{t+1} = \alpha P x_t + g$ (also known as value iteration in the context of DP), since the mapping $x \mapsto \alpha P x + g$ is a contraction with respect to the sup norm; see textbooks on DP, such as for example [6], or [20].

We focus on the case where n is very large, and it may be worth (even imperative) considering a low-dimensional approximation of a solution within a subspace

$$S = \{\Phi r \mid r \in \Re^s\},$$

where the columns of the $n \times s$ matrix Φ can be viewed as basis functions. This type of approximation approach has been the subject of much recent research in approximate DP, where several methods have been proposed and substantial computational experience has been accumulated. The most popular of these methods use projection with respect to the weighted Euclidean norm given by

$$\|x\|_\xi = \sqrt{\sum_{i=1}^n \xi_i x_i^2},$$

where $\xi \in \Re^n$ is a probability distribution with positive components. We denote by Π the projection operation onto S with respect to this norm (while Π depends on ξ, we do not show the dependence, since the associated vector ξ will always be clear from the context). The aforementioned methods for approximating the solution of the DP equation $x = \alpha P x + g$ aim to solve the equation

$$\Phi r = \Pi(\alpha P \Phi r + b)$$

with ξ being the invariant distribution of the transition probability matrix P (which is assumed irreducible; i.e., has a single recurrent class and no transient states). The more general methods of this paper aim to approximate a fixed point of the mapping

$$T(x) = Ax + b,$$

by solving the equation

$$\Phi r = \Pi T(\Phi r) = \Pi(A\Phi r + b), \tag{1.2}$$

where the projection norm $\|\cdot\|_\xi$ is determined in part by the structure of A in a way to induce some desired property. We view Π as a matrix and *we implicitly assume throughout that $I - \Pi A$ is invertible*. Thus, for a given ξ, there is a unique vector y^* such that $y^* = \Pi T(y^*)$, and we have $y^* = \Phi r^*$ for some $r^* \in \Re^s$ (if Φ has linearly independent columns, r^* is also unique).

To evaluate the distance between Φr^* and a fixed point x^* of T, we write

$$x^* - \Phi r^* = x^* - \Pi x^* + \Pi x^* - \Phi r^* = x^* - \Pi x^* + \Pi T x^* - \Pi T \Phi r^* = x^* - \Pi x^* + \Pi A(x^* - \Phi r^*), \tag{1.3}$$

from which

$$x^* - \Phi r^* = (I - \Pi A)^{-1}(x^* - \Pi x^*).$$

Thus, we have for any norm $\|\cdot\|$ and fixed point x^* of T

$$\|x^* - \Phi r^*\| \le \|(I - \Pi A)^{-1}\| \, \|x^* - \Pi x^*\|, \tag{1.4}$$

and the approximation error $\|x^* - \Phi r^*\|$ is proportional to the distance of the solution x^* from the approximation subspace. If ΠT is a contraction mapping of modulus $\alpha \in (0, 1)$ with respect to $\|\cdot\|$, from Eq. (1.3), we have

$$\|x^* - \Phi r^*\| \le \|x^* - \Pi x^*\| + \|\Pi T(x^*) - \Pi T(\Phi r^*)\| \le \|x^* - \Pi x^*\| + \alpha \|x^* - \Phi r^*\|,$$

so that

$$\|x^* - \Phi r^*\| \le \frac{1}{1-\alpha} \|x^* - \Pi x^*\|. \tag{1.5}$$

A better bound is obtained when ΠT is a contraction mapping of modulus $\alpha \in [0, 1)$ with respect to a Euclidean norm (e.g., $\|\cdot\|_\xi$). Then, using the Pythagorean Theorem, we have

$$\begin{aligned} \|x^* - \Phi r^*\|^2 &= \|x^* - \Pi x^*\|^2 + \|\Pi x^* - \Phi r^*\|^2 \\ &= \|x^* - \Pi x^*\|^2 + \|\Pi T(x^*) - \Pi T(\Phi r^*)\|^2 \\ &\le \|x^* - \Pi x^*\|^2 + \alpha^2 \|x^* - \Phi r^*\|^2 \end{aligned}$$

from which we obtain

$$\|x^* - \Phi r^*\|^2 \le \frac{1}{1-\alpha^2} \|x^* - \Pi x^*\|^2. \tag{1.6}$$