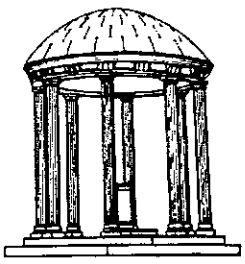
On Accelerating Monte Carlo Techniques for Solving Large Systems of Equations

TR96-041
1996

John H. Halton

Department of Computer Science
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599-3175



ON ACCELERATING MONTE CARLO TECHNIQUES FOR SOLVING LARGE SYSTEMS OF EQUATIONS

John H. Halton
The University of North Carolina at Chapel Hill

1. INTRODUCTION

This paper is concerned with ways of incorporating current Monte Carlo techniques for solving large linear systems [hereinafter referred-to as "plain Monte Carlo"—PMC] in accelerative schemes and other numerical techniques for more rapidly solving both linear and nonlinear systems.

Let $\mathbb{S}$ be a system of equations, linear or nonlinear, whose solution is an m-dimensional vector $\mathbf{x}$. Write

$$\mathbf{x} = \mathbf{y} + \mathbf{z}, \tag{1.1}$$

where $\mathbf{y}$ is an *estimate* of $\mathbf{x}$ and $\mathbf{z}$ is the corresponding *correction*. Call

$$\delta = \|\mathbf{x} - \mathbf{y}\|_\infty = \|\mathbf{z}\|_\infty \tag{1.2}$$

the *error*[1] in the estimate $\mathbf{y}$. Let

$$\mathbb{L} = \mathbb{L}(\mathbb{S}, \mathbf{y}) \tag{1.3}$$

be a linear problem, based on $\mathbb{S}$ and the estimate $\mathbf{y}$, whose solution is $\mathbf{z}$. Let

$$\mathcal{M} = \mathcal{M}(\mathbb{L}, \mathbf{Z}) \tag{1.4}$$

be an algorithm which generates and solves $\mathbb{L}$ to yield an estimate $\mathbf{Z}$ of $\mathbf{z}$. We can then take $\mathbf{y} + \mathbf{Z}$ as an improved estimate of $\mathbf{x}$. While the algorithm $\mathcal{M}$ may be deterministic or stochastic, we shall think particularly of the case

$$\mathcal{M} \equiv \text{PMC}. \tag{1.5}$$

Write

$$\Delta = \|\mathbf{z} - \mathbf{Z}\|_\infty \tag{1.6}$$

for the error in $\mathbf{Z}$.

Now suppose that we begin with the problem $\mathbb{S}$ and an initial

---

[1] Here we choose to have the *norm* of any m-dimensional vector $\mathbf{v} = (v_1, v_2, \ldots, v_m)$ to be the *maximum* (or L^∞) *norm*, $\|\mathbf{v}\|_\infty = \max_{1 \le j \le m} |v_j|$.

estimate $\mathbf{y}^{(0)}$ of its solution $\mathbf{x}$, and we set out to solve $\mathfrak{S}$ iteratively, by successively using the family of algorithms

$$\mathscr{M}_r = \mathscr{M}(\mathfrak{L}_r, \mathbf{Z}^{(r)}) \tag{1.7}$$

to generate and solve the corresponding linear problems

$$\mathfrak{L}_r = \mathfrak{L}(\mathfrak{S}, \mathbf{y}^{(r)}). \tag{1.8}$$

At each iteration,[2] $$\mathbf{x} = \mathbf{y}^{(r)} + \mathbf{z}^{(r)} \tag{1.9}$$

with[3] $$\delta_r = \|\mathbf{x} - \mathbf{y}^{(r)}\|_\infty = \|\mathbf{z}^{(r)}\|_\infty, \tag{1.10}$$

and we put $$\mathbf{y}^{(r+1)} = \mathbf{y}^{(r)} + \mathbf{Z}^{(r)}, \tag{1.11}$$

with[4] $$\Delta_r = \|\mathbf{z}^{(r)} - \mathbf{Z}^{(r)}\|_\infty, \tag{1.12}$$

and cycle through (1.7) and (1.8) until the error in the estimate $\mathbf{y}^{(r)}$,

$$\delta_r = \|\mathbf{z}^{(r)}\|_\infty < \varepsilon. \tag{1.13}$$

Clearly, by (1.9) and (1.11),

$$\mathbf{z}^{(r+1)} = \mathbf{x} - \mathbf{y}^{(r+1)} = \mathbf{x} - \mathbf{y}^{(r)} - \mathbf{Z}^{(r)} = \mathbf{z}^{(r)} - \mathbf{Z}^{(r)}, \tag{1.14}$$

whence, by (1.10) and (1.12),

$$\delta_{r+1} = \|\mathbf{z}^{(r+1)}\|_\infty = \|\mathbf{z}^{(r)} - \mathbf{Z}^{(r)}\|_\infty = \Delta_r. \tag{1.15}$$

---

2 Compare (1.1).

3 See (1.2).

4 See (1.6).

2. SEQUENTIAL MONTE CARLO FOR LINEAR SYSTEMS

One example of this kind of process is the *sequential Monte Carlo*—SMC—method, in which the problem $\mathbb{S}$ is itself linear, of the form

$$\mathbf{x} = \mathbf{a} + \mathbf{H}\mathbf{x}, \tag{2.1}$$

the initial iterate is

$$\mathbf{y}^{(0)} = \mathbf{0}, \tag{2.2}$$

so that, by (1.9),

$$\mathbf{x} = \mathbf{z}^{(0)}, \tag{2.3}$$

and the initial linear problem $\mathbb{L}_0 = \mathbb{L}(\mathbb{S}, \mathbf{y}^{(0)})$ is cast in the form

$$\mathbf{z}^{(0)} = \mathbf{d}^{(0)} + \mathbf{H}\mathbf{z}^{(0)}, \tag{2.4}$$

with

$$\mathbf{d}^{(0)} = \mathbf{a}. \tag{2.5}$$

Now PMC (i.e., Algorithm $\mathcal{M}_0 = \mathcal{M}(\mathbb{L}_0, \mathbf{Z}^{(0)})$[5] is applied to (2.4), with a Markov probability matrix $\mathbf{P}$ and a scoring scheme Σ (these are generally chosen once and for all), to yield a stochastic estimate $\mathbf{Z}^{(0)}$ of $\mathbf{z}^{(0)}$, with error $\Delta_0 = \delta_1$ which is usually estimated by the sample standard deviation [ssd] σ_0 of the scores whose average is $\mathbf{Z}^{(0)}$. We then obtain the new iterate

$$\mathbf{y}^{(1)} = \mathbf{y}^{(0)} + \mathbf{Z}^{(0)} = \mathbf{Z}^{(0)}. \tag{2.6}$$

For each $r \geq 1$, in the same way, Algorithm $\mathcal{M}_r = \mathcal{M}(\mathbb{L}_r, \mathbf{Z}^{(r)})$ then computes

$$\mathbf{d}^{(r)} = \mathbf{a} + \mathbf{H}\mathbf{y}^{(r)} - \mathbf{y}^{(r)}, \tag{2.7}$$

whence, by (1.9) and (2.1),

$$\mathbf{d}^{(r)} = \mathbf{a} + (\mathbf{H}\mathbf{x} - \mathbf{H}\mathbf{z}^{(r)}) - (\mathbf{x} - \mathbf{z}^{(r)}) = \mathbf{z}^{(r)} - \mathbf{H}\mathbf{z}^{(r)}$$

or[6]

$$\mathbf{z}^{(r)} = \mathbf{d}^{(r)} + \mathbf{H}\mathbf{z}^{(r)}. \tag{2.8}$$

---

5 See (1.4), (1.5), and (1.8).

6 Compare (2.4).

Algorithm $\mathscr{M}_r$ solves this problem by PMC, yielding an estimate $\mathbf{Z}^{(r)}$, and we accumulate

$$\mathbf{y}^{(r)} = \sum_{q=0}^{r} \mathbf{Z}^{(q)}. \tag{2.9}$$

It is then known that, if the problem is intrinsically convergent, a condition ensured by making

$$\rho(\mathbf{H}) < 1, \quad \rho(\mathbf{H}^+) < 1, \quad \rho(\mathbf{K}) < 1, \tag{2.10}$$

where $\rho(\mathbf{A})$ denotes the spectral radius of a matrix $\mathbf{A}$,[7]

$$(\mathbf{H}^+)_{ij} = |H_{ij}| \quad \text{and} \quad (\mathbf{K})_{ij} = \frac{H_{ij}^2}{P_{ij}}; \tag{2.11}$$

and, further, if the number of scores computed in each sequential iteration is sufficiently large,[8] then the error Δ_r in the stochastic estimate $\mathbf{Z}^{(r)}$, measured by its ssd, σ_r, decreases *geometrically* [*linearly*, *exponentially*]; i.e., there are constants Ξ and ξ, such that $|\xi| < 1$ and

$$\sigma_r \sim \Xi\xi^r \quad \text{as} \quad r \to \infty. \tag{2.12}$$

3. SEQUENTIAL MONTE CARLO FOR NONLINEAR SYSTEMS

Another example is a more complicated application of SMC, for a nonlinear problem,

$$\mathbf{F}(\mathbf{v}) = \mathbf{0}, \tag{3.1}$$

using Newton's method. By Taylor's Theorem, if we write

$$\mathbf{v} = \mathbf{w}^{(n)} + \mathbf{x}^{(n)}, \tag{3.2}$$

---

7 This is defined as the supremum of all absolute values [moduli] of eigenvalues:

$$\max\{\,|\lambda|:\ (\exists \mathbf{x} \neq \mathbf{0})\ \mathbf{H}\mathbf{x} = \lambda\mathbf{x}\,\}.$$

8 This number can be the *same* for all sequential iterations (stages).

then
$$\begin{aligned}\mathbf{F}(\mathbf{v}) &= \mathbf{F}(\mathbf{w}^{(n)} + \mathbf{x}^{(n)}) \\ &= \mathbf{F}(\mathbf{w}^{(n)}) + (\mathbf{x}^{(n)}.\nabla)\mathbf{F}(\mathbf{w}^{(n)}) + \frac{1}{2}(\mathbf{x}^{(n)}.\nabla)^2\mathbf{F}(\mathbf{w}^{(n)}) \\ &\quad + \frac{1}{6}(\mathbf{x}^{(n)}.\nabla)^3\mathbf{F}(\mathbf{w}^{(n)}) + \ldots = \mathbf{0}. \qquad (3.3)\end{aligned}$$

We linearize this problem to yield the approximation

$$\mathbf{F}(\mathbf{w}^{(n)}) + (\mathbf{x}^{(n)}.\nabla)\mathbf{F}(\mathbf{w}^{(n)}) = \mathbf{0}, \qquad (3.4)$$

i.e.,
$$\mathbf{J}(\mathbf{w}^{(n)})\mathbf{x}^{(n)} = -\mathbf{F}(\mathbf{w}^{(n)}), \qquad (3.5)$$

where $\mathbf{J}(\mathbf{w}^{(n)})$ is the value of the Jacobian matrix of $\mathbf{F}$ at $\mathbf{w}^{(n)}$:

$$(\mathbf{J}(\mathbf{w}))_{ij} = \frac{\partial F_i}{\partial w_j}. \qquad (3.6)$$

Now, we select an invertible (regular, non-singular) matrix $\mathbf{G}^{(n)}$ and put

$$\mathbf{a}^{(n)} = -\mathbf{G}^{(n)}\mathbf{F}(\mathbf{w}^{(n)}) \qquad (3.7)$$

and
$$\mathbf{H}^{(n)} = \mathbf{I} - \mathbf{G}^{(n)}\mathbf{J}(\mathbf{w}^{(n)}), \qquad (3.8)$$

yielding, by (3.5), that

$$\begin{aligned}\mathbf{H}^{(n)}\mathbf{x}^{(n)} &= \mathbf{x}^{(n)} - \mathbf{G}^{(n)}\mathbf{J}(\mathbf{w}^{(n)})\mathbf{x}^{(n)} \\ &= \mathbf{x}^{(n)} + \mathbf{G}^{(n)}\mathbf{F}(\mathbf{w}^{(n)}) = \mathbf{x}^{(n)} - \mathbf{a}^{(n)},\end{aligned}$$

i.e.,
$$\mathbf{x}^{(n)} = \mathbf{a}^{(n)} + \mathbf{H}^{(n)}\mathbf{x}^{(n)}. \qquad (3.9)$$

The analogy to (2.1) is immediate. We now apply SMC, as in §2,[9] to the solution of (3.9). This entails successive sequential stages, yielding stochastic estimates $\mathbf{Z}^{(n,r)}$ of $\mathbf{z}^{(n)}$ with ssd $\sigma^{(n,r)}$. As in §2, it is then known that the SMC method is intrinsically convergent[10] if, for all n,

$$\rho(\mathbf{H}^{(n)}) < 1, \quad \rho(\mathbf{H}^{(n)+}) < 1, \quad \rho(\mathbf{K}^{(n)}) < 1, \qquad (3.10)$$

and if the number of scores computed in each sequential iteration is sufficiently large. We then know that, if $0 < \alpha < 1$, and we continue

---

9 This is, in essence, (2.2)–(2.9), with the added superscript $^{(n)}$. Note: Algorithm $\mathscr{M}_r$ is now the entire SMC algorithm of §2, not PMC.

10 See (2.12).

the n-th Newtonian iteration's SMC until

$$\sigma^{(n,r)} < \alpha \frac{\|\mathbf{Z}^{(n)}\|_\infty^{\,3}}{\|\mathbf{Z}^{(n-1)}\|_\infty^{\,2}}, \tag{3.11}$$

we can find constants Ξ and ξ, such that $|\xi| < 1$ and

$$\|\mathbf{z}^{(n)}\|_\infty \sim \Xi\xi^{2^n} \quad \text{as} \quad r \to \infty; \tag{3.12}$$

i.e., the convergence is *quadratic*, as in Newton's method.

4. THE EIGENVALUE PROBLEM

The eigenvalue equation is

$$\mathbf{Hx} = \lambda\mathbf{x}, \tag{4.1}$$

with

$$\mathbf{x} \neq \mathbf{0}. \tag{4.2}$$

If $\mathbf{x}$ is a solution (called an *eigenvector*), so is any multiple $\kappa\mathbf{x}$, so long as κ is not zero. Thus, $\mathbf{x}$ really identifies an *eigendirection*. The eigenvector $\mathbf{x}$ and the *eigenvalue* λ are then said to belong to each other. If eigenvalues λ and μ both belong to the same eigenvector $\mathbf{x}$, we see by (4.1) and (4.2) that

$$\mathbf{Hx} = \lambda\mathbf{x} = \mu\mathbf{x} \quad \text{and} \quad \mathbf{x} \neq \mathbf{0};$$

so that

$$\lambda = \mu, \tag{4.3}$$

i.e., each eigenvector belongs to only one eigenvalue. Similarly, if eigenvectors $\mathbf{x}$ and $\mathbf{y}$ both belong to the same eigenvalue λ, then

$$\mathbf{Hx} = \lambda\mathbf{x} \quad \text{and} \quad \mathbf{Hy} = \lambda\mathbf{y},$$

whence, for any α and β,

$$\mathbf{H}(\alpha\mathbf{x} + \beta\mathbf{y}) = \lambda(\alpha\mathbf{x} + \beta\mathbf{y}); \tag{4.4}$$

More generally, we see that there is always an entire *eigensubspace* belonging to any given eigenvalue, and all such eigensubspaces are disjoint [as is usual in vector space theory, we ignore the null vector $\mathbf{0}$, which is in every subspace, but is *not* an eigenvector].

As is well known, the eigenvalues of a given $(m \times m)$ matrix $\mathbf{H}$ are solutions of the polynomial equation of degree m,

$$\det(\mathbf{H} - \lambda\mathbf{I}) = 0, \tag{4.5}$$

which is called the *characteristic equation*. It has just m solutions $\lambda_1, \lambda_2, \ldots, \lambda_m$ (if we take multiplicity into account), which can always be ordered so that

$$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_m|. \tag{4.6}$$

In the general case, the set of all eigenvectors can (if necessary) be augmented by further vectors, to form a base of the m-dimensional vector space. In this paper, we shall assume[11] that

$$|\lambda_1| > |\lambda_2| > |\lambda_3| > \ldots > |\lambda_m| \geq 0, \tag{4.7}$$

and then corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ are *linearly independent*, so that any vector $\mathbf{v}$ can be written uniquely as

$$\mathbf{v} = \xi_1\mathbf{x}_1 + \xi_2\mathbf{x}_2 + \xi_3\mathbf{x}_3 + \ldots + \xi_m\mathbf{x}_m. \tag{4.8}$$

Suppose that we have a $\mathbf{v}$ such that

$$\xi_1 \neq 0 \quad \text{and } \xi_2 \neq 0. \tag{4.9}$$

[Statistically, this is *almost surely* the case.] Then

$$\mathbf{Hv} = \xi_1\lambda_1\mathbf{x}_1 + \xi_2\lambda_2\mathbf{x}_2 + \ldots + \xi_m\lambda_m\mathbf{x}_m. \tag{4.10}$$

---

11 The complications arising if eigenvalues are not distinct are well-understood and their treatment, while not easy, is described in the literature; appropriate techniques are available. In a statistical sense, it is highly unlikely that any two eigenvalues should be equal, or should even have the same magnitude, unless the particular class of problems considered constrains such equality.

5. THE POWER METHOD

Write $\mathbf{O}(f(n))$ for an asymptotic characterization of a *vector* whose *norm* increases with $n \to \infty$ no faster than a given (usually, a simple) function $f(n)$.[12] Then we see, by (4.8) and (4.9), that we can write

$$\mathbf{u}^{(0)} = \mathbf{v} \tag{5.1}$$

and, as $r \to \infty$,

$$\mathbf{u}^{(r)} = \mathbf{H}^r\mathbf{v} = \xi_1\lambda_1{}^r\mathbf{x}_1 + \xi_2\lambda_2{}^r\mathbf{x}_2 + \mathbf{O}(|\lambda_3|^r). \tag{5.2}$$

If we put $$\kappa = \xi_2/\xi_1, \quad \alpha = \lambda_2/\lambda_1, \quad \text{and} \quad \beta = \lambda_3/\lambda_1, \tag{5.3}$$

so that, by (4.7), $$|\beta| < |\alpha| < 1, \tag{5.4}$$

we see, by (5.2), that, as $r \to \infty$,

$$\mathbf{u}^{(r)} = \xi_1\lambda_1{}^r\{\mathbf{x}_1 + \kappa\alpha^r\mathbf{x}_2 + \mathbf{O}(|\beta|^r)\}, \tag{5.5}$$

or $$\mathbf{u}^{(r)} = \xi_1\lambda_1{}^r\{\mathbf{x}_1 + \mathbf{O}(|\alpha|^r)\}. \tag{5.6}$$

In particular, $$\mathbf{u}^{(r)} \sim \xi_1\lambda_1{}^r\mathbf{x}_1; \tag{5.7}$$

i.e., for all i, $$u_i^{(r)} \sim \xi_1\lambda_1{}^r x_{1i}. \tag{5.8}$$

Since an eigenvector really identifies an *eigendirection*,[13] we can always assume, without loss of generality, that the base vectors are all *normalized*, each having at least one maximal component having value 1, i.e.,

$$\|\mathbf{x}_1\|_\infty = \|\mathbf{x}_2\|_\infty = \|\mathbf{x}_3\|_\infty = \ldots = \|\mathbf{x}_m\|_\infty = 1, \tag{5.9}$$

with $$x_{1h_1} = x_{2h_2} = x_{3h_3} = \ldots = x_{mh_m} = 1. \tag{5.10}$$

Of course, this defines the indices $h_1, h_2, \ldots, h_m$, to within a possible

---

12 This is an extension of the well-known "big-Oh" notation for asymptotics: the boldface **O** indicates a vector.

13 See the explanation just after (3.10).

(unimportant) amount of ambiguity [if several components x_{ih_i} have magnitude $|x_{ih_i}| = 1$]. By (4.7), (4.9), (5.8), and (5.10), we observe that, for all sufficiently large r,

$$u_{h_1}^{(r)} \neq 0. \tag{5.11}$$

Now let us *scale* the vectors $\mathbf{u}^{(r)}$ defined above, by the relation

$$\mathbf{z}^{(0)} = (1/\sigma_0)\mathbf{u}^{(0)} = (1/\sigma_0)\mathbf{v} \tag{5.12}$$

and

$$\mathbf{z}^{(r)} = (1/\sigma_r)\mathbf{u}^{(r)}. \tag{5.13}$$

If we put

$$\mathbf{z}^{(r)} = (1/\tau_r)\mathbf{H}\mathbf{z}^{(r-1)}, \tag{5.14}$$

then, by (5.1), (5.8), and (5.14),

$$\begin{aligned}\mathbf{z}^{(r)} &= (1/\sigma_r)\mathbf{u}^{(r)} = (1/\tau_r)\mathbf{H}\mathbf{z}^{(r-1)} = (1/\tau_r\tau_{r-1})\mathbf{H}^2\mathbf{z}^{(r-2)} \\ &= \ldots = (1/\tau_r\,\tau_{r-1}\ldots\tau_1)\mathbf{H}^r\mathbf{z}^{(0)} = (1/\tau_r\,\tau_{r-1}\ldots\tau_1\sigma_0)\mathbf{H}^r\mathbf{v} \\ &= (1/\tau_r\,\tau_{r-1}\ldots\tau_1\sigma_0)\mathbf{u}^{(r)},\end{aligned}$$

whence

$$\sigma_r = \tau_r\,\tau_{r-1}\ldots\tau_1\sigma_0, \tag{5.15}$$

and

$$\tau_r = \sigma_r/\sigma_{r-1}. \tag{5.16}$$

If we now specify that the vectors $\mathbf{z}^{(r)}$ be *normalized*, too—

$$\|\mathbf{z}^{(r)}\|_\infty = 1, \tag{5.17}$$

then, by (5.8),

$$|\sigma_r| = \|\mathbf{u}^{(r)}\|_\infty. \tag{5.18}$$

We note that the σ_r are not yet known beyond their absolute values [moduli], so that we can choose their sign (or phase angle, if they are complex) arbitrarily. As for the base vectors, we assume that at least one maximal component of $\mathbf{z}^{(r)}$ has value 1. Let this component[14] be $z_{k_r}^{(r)}$. Then, by (5.4), (5.5), (5.11), and (5.8),[15]

$$\mathbf{z}^{(r)} \to \mathbf{x}_1 \quad \text{as} \quad r \to \infty \tag{5.19}$$

---

14 This is modulo the ambiguity referred-to just after (5.10).

15 In the case of any ambiguity, we try to keep the index k_r constant, rather than allow it to fluctuate unnecessarily.

and, for all sufficiently large r, we can put

$$z_{h_1}^{(r)} = 1; \tag{5.20}$$

i.e.,

$$k_r = h_1. \tag{5.21}$$

Furthermore, for these values of r, by (5.14),

$$\tau_r = (\mathbf{Hz}^{(r-1)})_{h_1}. \tag{5.22}$$

Also,

$$\sigma_r = u_{h_1}^{(r)}/z_{h_1}^{(r)} = \xi_1\lambda_1{}^r\{1 + \kappa\alpha^r x_{2h_1} + O(|\beta|^r)\}, \tag{5.23}$$

so

$$\sigma_r = \xi_1\lambda_1{}^r\{1 + O(|\alpha|^r)\}. \tag{5.24}$$

Therefore, by (4.9), (5.4), (5.13), (5.16), and (5.23), as $r \to \infty$,[16]

$$\begin{aligned}\tau_r &= (\mathbf{Hz}^{(r-1)})_{h_1} = (1/\sigma_{r-1})\,(\mathbf{Hu}^{(r-1)})_{h_1} \\ &= \frac{\xi_1\lambda_1{}^r[x_{1h_1} + \kappa\alpha^r x_{2h_1} + O(|\beta|^r)]}{\xi_1\lambda_1{}^{r-1}[1 + \kappa\alpha^{r-1}x_{2h_1} + O(|\beta|^{r-1})]} \\ &= \lambda_1\{1 - \kappa(1-\alpha)\alpha^{r-1}x_{2h_1} + O(|\beta|^r)\} = \lambda_1\{1 + O(|\alpha|^r)\}.\end{aligned} \tag{5.25}$$

Thus

$$\tau_r \to \lambda_1 \quad \text{as} \quad r \to \infty. \tag{5.26}$$

Furthermore, the convergence is *geometric*, since the error is given by (5.4) and (5.25) as $O(|\alpha|^r)$. This is called the *Power Method*.

---

16 Here, we use the well-known properties of asymptotic expressions, that

(i) $O(|\alpha|^{r-1}) = O(|\alpha|^r)$,

and (ii) $\{1 + O(|\alpha|^r)\}/\{1 + O(|\alpha|^r)\} = 1 + O(|\alpha|^r)$.

6. RAYLEIGH QUOTIENTS

If the matrix $\mathbf{H}$ has additional properties, even faster methods are available. For example, if $\mathbf{H}$ is Hermitian, i.e.,[17]

$$\mathbf{H} = \mathbf{H}^* \quad \text{or} \quad (\forall i,j)\ H_{ij} = H_{ji}^*, \tag{6.1}$$

then $\mathbf{x}_i^*\mathbf{H}\mathbf{x}_i = \mathbf{x}_i^*(\mathbf{H}\mathbf{x}_i) = \mathbf{x}_i^*(\lambda_i\mathbf{x}_i) = \lambda_i\mathbf{x}_i^*\mathbf{x}_i$

$$= (\mathbf{x}_i^*\mathbf{H})\mathbf{x}_i = (\mathbf{x}_i^*\mathbf{H}^*)\mathbf{x}_i = (\mathbf{H}\mathbf{x}_i)^*\mathbf{x}_i = (\lambda_i\mathbf{x}_i)^*\mathbf{x}_i = \lambda_i^*\mathbf{x}_i^*\mathbf{x}_i,$$

whence
$$\lambda_i^* = \lambda_i; \tag{6.2}$$

that is, all the eigenvalues are *real*; and if $i \neq j$ and $\lambda_i \neq \lambda_j$,

$$\mathbf{x}_i^*\mathbf{H}\mathbf{x}_j = \mathbf{x}_i^*(\mathbf{H}\mathbf{x}_j) = \mathbf{x}_i^*\lambda_j\mathbf{x}_j = \lambda_j\mathbf{x}_i^*\mathbf{x}_j$$

$$= (\mathbf{x}_i^*\mathbf{H})\mathbf{x}_j = (\mathbf{H}\mathbf{x}_i)^*\mathbf{x}_j = (\lambda_i\mathbf{x}_i)^*\mathbf{x}_j = \lambda_i\mathbf{x}_i^*\mathbf{x}_j;$$

so that
$$\mathbf{x}_i^*\mathbf{x}_j = 0, \tag{6.3}$$

i.e., the eigenvectors belonging to distinct eigenvalues are *orthogonal*. Hence, by (5.5), (5.8), and (6.3), with a little simplification, we get that

$$\mathfrak{R}_r = (\mathbf{H}\mathbf{z}^{(r)})^*\mathbf{z}^{(r)} / \mathbf{z}^{(r)*}\mathbf{z}^{(r)} = (\mathbf{H}\mathbf{u}^{(r)})^*\mathbf{u}^{(r)} / \mathbf{u}^{(r)*}\mathbf{u}^{(r)}$$

$$= \frac{\xi_1\lambda_1^{r+1}\{\mathbf{x}_1^* + \kappa(\alpha^*)^{r+1}\mathbf{x}_2^* + \mathbf{O}(|\beta|^r)\}\{\mathbf{x}_1 + \kappa\alpha^r\mathbf{x}_2 + \mathbf{O}(|\beta|^r)\}}{\xi_1\lambda_1^{r}\{\mathbf{x}_1^* + \kappa(\alpha^*)^{r}\mathbf{x}_2^* + \mathbf{O}(|\beta|^r)\}\{\mathbf{x}_1 + \kappa\alpha^r\mathbf{x}_2 + \mathbf{O}(|\beta|^r)\}}$$

$$= \lambda_1 \frac{\mathbf{x}_1^*\mathbf{x}_1 + \kappa|\alpha|^{2r}\mathbf{x}_2^*\mathbf{x}_2 + O(|\beta|^{2r})}{\mathbf{x}_1^*\mathbf{x}_1 + \kappa|\alpha|^{2r}\mathbf{x}_2^*\mathbf{x}_2 + O(|\beta|^{2r})}$$

$$= \lambda_1 \frac{1 + O(|\alpha|^{2r})}{1 + O(|\alpha|^{2r})} = \lambda_1[1 + O(|\alpha|^{2r})]. \tag{6.4}$$

Thus,
$$\mathfrak{R}_r \to \lambda_1 \quad \text{as} \quad r \to \infty; \tag{6.5}$$

---

17 For a possibly complex number $z = x + iy$, where x and y are real numbers, z^* denotes the *complex conjugate* number, $z^* = x - iy$. For a possibly complex matrix, $\mathbf{H} = \mathbf{L} + i\mathbf{M}$, again with matrices $\mathbf{L}$ and $\mathbf{M}$ real, $\mathbf{H}^*$ denotes the *Hermitian transpose*, $\mathbf{H}^* = \mathbf{L}^T - i\mathbf{M}^T$, as indicated in (6.1).

and, furthermore, the convergence is *twice as fast* as for the regular Power Method. $\mathfrak{R}_r$ is called the *Rayleigh quotient*.

7. THE ITERATE DIFFERENCE

Now consider the *difference* $\mathbf{z}^{(r)} - \mathbf{z}^{(r-1)}$. By (5.4), (5.13), (5.16), (5.14), and (6.1),

$$\begin{aligned}
\mathbf{z}^{(r)} - \mathbf{z}^{(r-1)} &= (1/\sigma_r)\mathbf{u}^{(r)} - (1/\sigma_{r-1})\mathbf{u}^{(r-1)} \\
&= (1/\sigma_r)\{\xi_1\lambda_1{}^r[\mathbf{x}_1 + \kappa\alpha^r\mathbf{x}_2 + \mathbf{O}(|\beta|^r)]\} \\
&\qquad - (1/\sigma_{r-1})\{\xi_1\lambda_1{}^{r-1}[\mathbf{x}_1 + \kappa\alpha^{r-1}\mathbf{x}_2 + \mathbf{O}(|\beta|^{r-1})]\} \\
&= (\xi_1\lambda_1{}^r/\sigma_r)\{(1 - \tau_r/\lambda_1)\mathbf{x}_1 + \kappa(1 - \tau_r/\lambda_2)\alpha^r\mathbf{x}_2 + \mathbf{O}(|\beta|^r)\} \\
&= \{1 + O(|\alpha|^r)\}\{(1 - \tau_r/\lambda_1)\mathbf{x}_1 + \kappa(1 - \tau_r/\lambda_2)\alpha^r\mathbf{x}_2 + \mathbf{O}(|\beta|^r)\}.
\end{aligned}$$

We observe, by (5.25), that

$$1 - \tau_r/\lambda_1 = \kappa(1 - \alpha)\alpha^{r-1}x_{2h_1} + \mathrm{O}(|\beta|^r) = \alpha^r O(1), \tag{7.1}$$

while

$$1 - \tau_r/\lambda_2 = O(1); \tag{7.2}$$

so that

$$\mathbf{z}^{(r)} - \mathbf{z}^{(r-1)} = \alpha^r\{1 + O(|\alpha|^r)\}\{A\mathbf{x}_1 + B\mathbf{x}_2 + \mathbf{O}(|\beta/\alpha|^r)\}, \tag{7.3}$$

where A and B are constants independent of r. Thus, we can estimate both λ_2 (through α) and $\mathbf{x}_2$ (since we presumably have estimates of λ_1 and $\mathbf{x}_1$) from the sequence of differences, much as we estimate the principal eigenvalue λ_1 and eigenvector $\mathbf{x}_1$ from the $\mathbf{z}^{(r)}$. Here, again, convergence is *geometric*.

8. THE INVERSE POWER METHOD

We return to the eigenvalue problem—equations (4.1) and (4.2)—discussed in §4–§7. We shall assume that all eigenvalues of the matrix $\mathbf{H}$ are of different magnitudes[18], and also that the matrix $\mathbf{H}$ is Hermitian,[19] so that its eigenvalues λ_i are real and its eigenvectors $\mathbf{x}_i$ form a base of m-space and are orthogonal, as in (6.3); indeed, we can assume that they are *orthonormal*:[20]

$$(\forall i, j) \quad \mathbf{x}_i^* \mathbf{x}_j = \delta_{ij}. \tag{8.1}$$

The theory presented earlier still applies, slightly modified by the new normalization. We see that (5.9) and (5.10) are replaced by

$$\|\mathbf{x}_1\|_2 = \|\mathbf{x}_2\|_2 = \|\mathbf{x}_3\|_2 = \ldots = \|\mathbf{x}_m\|_2 = 1, \tag{8.2}$$

and (5.17) and (5.18) are replaced by

$$\|\mathbf{z}^{(r)}\|_2 = 1 \tag{8.3}$$

and

$$|\sigma_r| = \|\mathbf{u}^{(r)}\|_2. \tag{8.4}$$

Thus, by (5.14) and (5.19) (which still holds),

$$\mathbf{x}_1 \sim (1/\tau_r)\mathbf{H}\mathbf{x}_1 = (1/\tau_r)\lambda_1 \mathbf{x}_1;$$

which implies (5.26), as before. Furthermore, the power method of §6 is entirely unchanged.

---

18 See (4.7).

19 See (6.1).

20 Here, δ_{ij} is "Kronecker's delta"—$\delta_{ij} = 1$ if $i = j$, $\delta_{ij} = 0$ if $i \neq j$.

This assumption replaces the normalization implied in (5.9) and (5.10). Instead of the L^∞ norm, $\|\mathbf{v}\|_\infty = \max_{1 \le j \le m} |v_j|$, we use the L^2 norm, $\|\mathbf{v}\|_2 = (\mathbf{v}^*\mathbf{v})^{1/2} = \left(\sum_{j=1}^{m} |v_j|^2 \right)^{1/2}$.

Now let μ be any real number and consider the matrix

$$\mathbf{M} = \mathbf{H} - \mu\mathbf{I}, \tag{8.5}$$

Clearly, for any eigenvector $\mathbf{x}_i$ of $\mathbf{H}$,

$$\mathbf{M}\mathbf{x}_i = (\mathbf{H} - \mu\mathbf{I})\mathbf{x}_i = (\lambda_i - \mu)\mathbf{x}_i \tag{8.6}$$

If we assume that $\mathbf{M}$ is invertible,[21] and write

$$\mathbf{N} = \mathbf{M}^{-1} = (\mathbf{H} - \mu\mathbf{I})^{-1}, \tag{8.7}$$

so that

$$\mathbf{M}\mathbf{N} = \mathbf{N}\mathbf{M} = \mathbf{I}, \tag{8.8}$$

then[22]

$$\mathbf{N}\mathbf{x}_i = (\mathbf{H} - \mu\mathbf{I})^{-1}\mathbf{x}_i = (\lambda_i - \mu)^{-1}\mathbf{x}_i. \tag{8.9}$$

Therefore, by (4.8) and (4.10),

$$\mathbf{M}\mathbf{v} = \xi_1(\lambda_1 - \mu)\mathbf{x}_1 + \xi_2(\lambda_2 - \mu)\mathbf{x}_2 + \ldots + \xi_m(\lambda_m - \mu)\mathbf{x}_m \tag{8.10}$$

and

$$\mathbf{N}\mathbf{v} = \xi_1(\lambda_1 - \mu)^{-1}\mathbf{x}_1 + \xi_2(\lambda_2 - \mu)^{-1}\mathbf{x}_2 + \ldots + \xi_m(\lambda_m - \mu)^{-1}\mathbf{x}_m. \tag{8.11}$$

That is to say, the vectors $\mathbf{x}_i$ are also eigenvectors of the matrices $\mathbf{M}$ and $\mathbf{N}$.

Now consider carrying out the power method with the matrix $\mathbf{N}$ instead of $\mathbf{H}$. Suppose that μ is nearest to the eigenvalue λ_s, and nearer to it than to any other eigenvalue, and write

$$\lambda_i = \mu + \nu_i. \tag{8.12}$$

Then

$$\min_{1\le i\le m} |\lambda_i - \mu| = |\lambda_s - \mu|, \tag{8.13}$$

or

$$\min_{1\le i\le m} |\nu_i| = |\nu_s|. \tag{8.14}$$

If we define the parameter κ by

$$0 < |\nu_s| < \kappa = \min_{i\ne s} |\lambda_i - \lambda_s|, \tag{8.15}$$

then

$$\min_{i\ne s} |\nu_i| = \min_{i\ne s} |\lambda_i - \mu| = \min_{i\ne s} |(\lambda_i - \lambda_s) + (\lambda_s - \mu)|,$$

---

21 This assumption is equivalent to assuming that μ is itself *not* an eigenvalue of $\mathbf{H}$.

22 To obtain (8.9), premultiply (8.6) by $\mathbf{N}$ and divide by the (presumed non-zero) scalar $\lambda_i - \mu$.

so that $$\min_{i \neq s} |\nu_i| \geq \kappa + |\nu_s|, \tag{8.16}$$

whence $$\max_{i \neq s} \{|\nu_i|^{-1}\} \leq (\kappa + |\nu_s|)^{-1} < \kappa^{-1}. \tag{8.17}$$

Let us also write $$0 < \alpha = \frac{|\nu_s|}{\kappa + |\nu_s|} < 1. \tag{8.18}$$

Arguing just as in §5, we see that, if[23]

$$\mathbf{u}^{(0)} = \mathbf{v} \quad \text{and} \quad \mathbf{u}^{(r)} = \mathbf{N}^r \mathbf{v}, \tag{8.19}$$

then, as $r \to \infty$,[24]

$$\mathbf{u}^{(r)} = \xi_s \nu_s^{-r} \mathbf{x}_s + \mathbf{O}((\kappa + |\nu_s|)^{-r}) = \xi_s \nu_s^{-r} [\mathbf{x}_s + \mathbf{O}(\alpha^r)]. \tag{8.20}$$

As before, we can normalize the $\mathbf{u}^{(r)}$ as $\mathbf{z}^{(r)}$with (8.3),[25] and then

$$\mathbf{u}^{(r)*}\mathbf{u}^{(r)} = |\xi_s \nu_s^{-r}|^2 [1 + O(\alpha^{2r})],$$

whence[26]

$$\|\mathbf{u}^{(r)}\|_2 = (\mathbf{u}^{(r)*}\mathbf{u}^{(r)})^{1/2} = |\xi_s \nu_s^{-r}| [1 + O(\alpha^{2r})], \tag{8.21}$$

and so[27]

$$\mathbf{z}^{(r)} = (\mathbf{u}^{(r)*}\mathbf{u}^{(r)})^{-1/2}\mathbf{u}^{(r)} = [1 + O(\alpha^{2r})] [\mathbf{x}_s + \mathbf{O}(\alpha^r)]. \tag{8.22}$$

Hence, as $r \to \infty$, $$\mathbf{z}^{(r)} \to \mathbf{x}_s. \tag{8.23}$$

---

23 See (5.1) and (5.2).

24 Compare (5.6).

25 See footnote [20].

26 Observe that, by the Binomial Theorem, since $|\alpha| < 1$, as $r \to \infty$,

$$[1 + O(\alpha^{2r})]^{1/2} = 1 + \frac{1}{2}O(\alpha^{2r}) - \frac{1}{8}O(\alpha^{4r}) + \ldots = 1 + O(\alpha^{2r}).$$

27 See footnote [16]. Furthermore, the sign of $\mathbf{x}_s$ is ambiguous. By keeping the sign of some component of $\mathbf{z}^{(r)}$, say $z_h^{(r)}$, constant, as $r \to \infty$, we can always make $\mathbf{z}^{(r)} \sim \mathbf{x}_s$, rather than $\mathbf{z}^{(r)} \sim -\mathbf{x}_s$.

If we define $$\mathbf{y}^{(r)} = \mathbf{N}\mathbf{z}^{(r-1)} = (\mathbf{u}^{(r-1)*}\mathbf{u}^{(r-1)})^{-1/2}\mathbf{u}^{(r)}, \tag{8.24}$$

then[28] $$\mathbb{T}_{r-1} = \frac{1}{\mathbf{y}^{(r)*}\mathbf{z}^{(r-1)}} = \frac{\mathbf{u}^{(r-1)*}\mathbf{u}^{(r-1)}}{\mathbf{u}^{(r)*}\mathbf{u}^{(r-1)}}$$

$$= \frac{\xi_s^{\,2}\nu_s^{-2r+2}\,[1 + O(\alpha^{2r})]}{\xi_s^{\,2}\nu_s^{-2r+1}\,[1 + O(\alpha^{2r})]} = \nu_s\,[1 + O(\alpha^{2r})]. \tag{8.25}$$

Thus, $$\mathbb{T}_r \to \nu_s \quad \text{as} \quad r \to \infty, \tag{8.26}$$

or, by (8.12), $$\mathbb{T}_r + \mu \to \lambda_s \quad \text{as} \quad r \to \infty. \tag{8.27}$$

This shows that this *inverse power method* converges to λ_s, the nearest eigenvalue to μ, as $r \to \infty$, and the convergence is again *geometric*.

The computational procedure is thus as follows.

0. Begin with an arbitrary vector $\mathbf{v} = \mathbf{u}^{(0)}$ and an arbitrary real number μ. Define $\mathbf{z}^{(0)} = (\mathbf{v}^*\mathbf{v})^{-1/2}\mathbf{v}$, so that $\|\mathbf{z}^{(0)}\|_2 = 1$. Take $r = 1$.

1. i. Define the matrix

$$\mathbf{M} = \mathbf{H} - \mu\mathbf{I}. \tag{8.28}$$

ii. Solve the equation

$$\mathbf{M}\mathbf{y}^{(r)} = \mathbf{z}^{(r-1)} \tag{8.29}$$

and put $$\mathbf{z}^{(r)} = (\mathbf{y}^{(r)*}\mathbf{y}^{(r)})^{-1/2}\mathbf{y}^{(r)}. \tag{8.30}$$

2. Compute $$\mathbb{T}_{r-1} = (\mathbf{y}^{(r)*}\mathbf{z}^{(r-1)})^{-1}. \tag{8.31}$$

3. Increment r. Repeat [1] and [2] until

$$\|\mathbf{z}^{(r)} - \mathbf{z}^{(r-1)}\|_2 < \varepsilon. \tag{8.32}$$

---

28 Compare (6.4), given (8.3).

9. ACCELERATING THE INVERSE POWER METHOD

We know that the inverse power method (described above) converges geometrically to λ_s. It is also clear that the rate of convergence of the method is governed by the constant α defined in (8.12)–(8.18). As ν_s decreases, so does α, and the convergence becomes faster. Thus, if we change μ so as to decrease ν_s in the course of the computation, we can only accelerate the convergence of the method to λ_s.

By (8.24) and (8.25), we can use $\mathfrak{T}_{r-1} = ((\mathbf{N}\mathbf{z}^{(r-1)})^*\mathbf{z}^{(r-1)})^{-1}$ $= (\mathbf{u}^{(r-1)*}\mathbf{u}^{(r-1)}/\mathbf{u}^{(r)*}\mathbf{u}^{(r-1)})^{-1}$ to approximate ν_s. Suppose, therefore, that we replace the Rayleigh quotient method outlined in §8 above by a two step sequential method as follows.[29]

0. Begin with an arbitrary vector $\mathbf{v} = \mathbf{u}^{(0)}$ and an arbitrary real number $\mu = \mu^{(0)}$. Define $\mathbf{z}^{(0)} = (\mathbf{v}^*\mathbf{v})^{-1/2}\mathbf{v}$, so that $\|\mathbf{z}^{(0)}\|_2 = 1$. Take $r = 1$.

1. i. Define the matrix

$$\mathbf{M}^{(r-1)} = \mathbf{H} - \mu^{(r-1)}\mathbf{I}. \tag{9.1}$$

ii. Solve the equation

$$\mathbf{M}^{(r-1)}\mathbf{y}^{(r)} = \mathbf{z}^{(r-1)} \tag{9.2}$$

and put

$$\mathbf{z}^{(r)} = (\mathbf{y}^{(r)*}\mathbf{y}^{(r)})^{-1/2}\mathbf{y}^{(r)}. \tag{9.3}$$

2. i. Compute $\mathfrak{T}_{r-1} = (\mathbf{y}^{(r)*}\mathbf{z}^{(r-1)})^{-1}$. (9.4)

ii. Put $\mu^{(r)} = \mu^{(r-1)} + \mathfrak{T}_{r-1}$. (9.5)

3. Increment r. Repeat [1] and [2] until

$$\|\mathbf{z}^{(r)} - \mathbf{z}^{(r-1)}\|_2 < \varepsilon. \tag{9.6}$$

---

29 Compare the closely similar procedure in (8.28)–(8.32). The initialization is [0]; the two major steps are [1] and [2]; the conditional looping command is [3]. The crucial change is the iterative improvement of $\mu^{(r)}$ through (9.5), in [2][ii].

The change from the inverse Rayleigh quotient method is solely in the line [2][ii]; in the original method, all $\mu^{(r)} = \mu^{(0)}$. Note, too, that the application of the *Monte Carlo* method is to perform [1][ii] by PMC.[30]

For the original method, we get, by (8.22), that

$$\mathbf{z}^{(r)} = [1 + O(\alpha^{2r})]\, \mathbf{x}_s + \mathbf{O}(\alpha^r), \tag{9.7}$$

whence
$$\mathbf{z}^{(r)} - \mathbf{x}_s = \mathbf{O}(\alpha^r); \tag{9.8}$$

i.e.,
$$\|\mathbf{z}^{(r)} - \mathbf{x}_s\|_2 = O(\alpha^r). \tag{9.9}$$

Similarly, by (8.25),

$$\left| \mathbb{T}_{r-1} - \nu_s \right| = \left| (\mathbb{T}_{r-1} + \mu) - \lambda_s \right| = O(\alpha^{2r}). \tag{9.10}$$

In the new algorithm, it is clear that ν_s must be replaced by a changing parameter,

$$\nu_s^{(r)} = \lambda_s - \mu^{(r)} = \lambda_s - \mu^{(r-1)} - \mathbb{T}_{r-1} = \nu_s^{(r-1)} - \mathbb{T}_{r-1}. \tag{9.11}$$

Going over the previous line of argument, we observe that, if we now write

$$P_r = \prod_{t=0}^{r-1} \alpha^{(t)} = \prod_{t=0}^{r-1} \frac{\left|\nu_s^{(t)}\right|}{\kappa} \geq \prod_{t=0}^{r-1} \frac{\left|\nu_s^{(t)}\right|}{\kappa + \left|\nu_s^{(t)}\right|}, \tag{9.12}$$

then[31]
$$\mathbf{z}^{(r)} = [1 + O(P_r^2)]\, \mathbf{x}_s + \mathbf{O}(P_r), \tag{9.13}$$

whence[32]
$$\mathbf{z}^{(r)} - \mathbf{x}_s = \mathbf{O}(P_r), \tag{9.14}$$

i.e.,[33]
$$\|\mathbf{z}^{(r)} - \mathbf{x}_s\|_2 = O(P_r). \tag{9.15}$$

---

30 It is also possible to replace the scalar products in (9.3) and (9.4) by MC estimates.

31 Compare (9.7).

32 Compare (9.8).

33 Compare (9.9).

The counterpart of (8.25) is now, by (9.4),

$$\mathbb{T}_{r-1} = v_s^{(r-1)} [1 + O(P_r^2)], \tag{9.16}$$

and so $$\left| \mathbb{T}_{r-1} - v_s^{(r-1)} \right| = \left| (\mathbb{T}_{r-1} + \mu^{(r-1)}) - \lambda_s \right| = O(P_r^2). \tag{9.17}$$

Note that, because each $\mu^{(r)}$ is an improvement on its predecessor, and therefore each $v_s^{(r)}$ is less than its predecessor, P_r will necessarily be less than the corresponding α^r in (9.7)–(9.10).

By (9.11) and (9.16), we see that

$$v_s^{(r)} = v_s^{(r-1)} - v_s^{(r-1)} [1 + O(P_r^2)] = v_s^{(r-1)} O(P_r^2). \tag{9.18}$$

We can attempt to solve the relations (9.12) and (9.18) by putting, for some

$$Q \geq 1, \quad C > 1, \quad \text{and} \quad 0 < q < 1, \tag{9.19}$$

that $$\alpha^{(t)} = \frac{\left| v_s^{(r)} \right|}{\kappa} \sim Q q^{C^r}. \tag{9.20}$$

Then, by (9.12), $$P_r = \prod_{t=0}^{r-1} \alpha^{(t)} \sim \prod_{t=0}^{r-1} Q q^{C^t}$$

$$= Q^r q^{1+C+C^2+\ldots+C^{r-1}} = Q^r q^{(C^r-1)/(C-1)}, \tag{9.21}$$

so that we need, by (9.18), that

$$q^{C^r - C^{r-1}} \sim \frac{\alpha^{(r)}}{\alpha^{(r-1)}} = O(Q^{2r} q^{2(C^r-1)/(C-1)}). \tag{9.22}$$

Since, by our assumption, $C > 1$ and $q < 1$, and (9.22) should hold for *all* sufficiently large r, we need

$$C^r - C^{r-1} \geq \frac{2(C^r - 1)}{C - 1} \geq \frac{2C^r}{C - 1}. \tag{9.23}$$

This holds for all r if $$(C-1)^2 \geq 2C,$$

i.e., if $$C^2 - 4C + 1 \geq 0,$$

i.e., if[34] $$C \geq 2 + \sqrt{3} \approx 3.732. \tag{9.24}$$

The tightest bound for (9.20) and (9.22) is clearly $C = 2 + \sqrt{3}$, i.e.,

$$\boxed{v_s^{(r)} \sim A\, q^{(2+\sqrt{3})^r} \approx A\, q^{3.732^r}}\ , \tag{9.25}$$

where $A = \kappa Q$. Hence, by (9.21),[35]

$$P_r \sim B\, Q^r\, q^{[(\sqrt{3}-1)/2]\,(2+\sqrt{3})^r} \approx B\, Q^r\, (q^{0.366})^{3.732^r}\ , \tag{9.26}$$

where $B = q^{-0.366}$. Of course, by (9.19),

$$0 < q < q^\dagger = q^{(\sqrt{3}-1)/2} \approx q^{0.366} < 1, \tag{9.27}$$

and, as $r \to \infty$, however large Q may be, $Q^r\, p^{\,(2+\sqrt{3})^r}$ tends to zero faster than $(q^\dagger + \varepsilon)^{\,(2+\sqrt{3})^r}$ for any $\varepsilon > 0$, however small. For relative simplicity, let us write, for some small ε,

$$p = q^\dagger + \varepsilon; \tag{9.28}$$

then, by (9.26), $$\boxed{P_r = o(p^{(2+\sqrt{3})^r}) \approx o(p^{3.732^r})}\ . \tag{9.29}$$

It now follows from (9.15) that

$$\boxed{\left\| \mathbf{z}^{(r)} - \mathbf{x}_s \right\|_2 = o(p^{(2+\sqrt{3})^r}) \approx o(p^{3.732^r})} \tag{9.30}$$

and from (9.17) that

$$\boxed{\left| \mathfrak{T}_{r-1} - v_s^{(r-1)} \right| = o(p^{2\times(2+\sqrt{3})^r}) = o((p^2)^{(2+\sqrt{3})^r}) \approx o((p^2)^{3.732^r})}\ . \tag{9.31}$$

This shows that the accelerative, sequential method increases the convergence of the process from *linear* (sometimes also called *geometric* or *exponential*) to *more-than-cubic*. By comparison, *Newton's method*, which has stood the test of time for over 300 years, converges only *quadratically*.

To illustrate this concept in a different way, suppose that we gain d

---

34 The full solution is $C \in (-\infty, 2 - \sqrt{3}] \cup [2 + \sqrt{3}, \infty)$; the former range is *inadmissible*, by our assumption that $C > 1$.

35 $(C - 1)^{-1} = (\sqrt{3} + 1)^{-1} = (\sqrt{3} - 1)/(3 - 1) = (\sqrt{3} - 1)/2 \approx 0.366$.

decimal places of accuracy by a single iteration of a linearly-converging method. Then this is because the error is multiplied by 10^{-d}. Thus, subsequent errors after this will be multiplied by 10^{-d} also; so that we shall gain d decimal places of accuracy at each iteration. However, if we were to gain d decimal places of accuracy by a single iteration of a quadratically-converging method, like Newton's method, then we would have, for some $T > 0$ and $0 < t < 1$,

$$10^{-d} = \frac{Tt^{2^{r+1}}}{Tt^{2^r}} = t^{2^r}, \tag{9.32}$$

and subsequent errors would be $Tt^{2^{r+2}}$, $Tt^{2^{r+3}}$, and so on, successively multiplying the current error, first, by $t^{2^{r+1}} = (t^{2^r})^2 = 10^{-2d}$, then by $t^{2^{r+2}} = (t^{2^r})^{2^2} = 10^{-4d}$, and so on, *doubling* the number of accurate decimal digits gained at each step. This is what leads to the spectacular convergence for which Newton's method is rightly admired. Finally, if we use the present method, it is easily seen that the number of accurate decimal digits gained at each step is successively *multiplied by about* 3.732!

Although no serious setbacks have been encountered in preliminary computational experiments, two notes of caution must be made. First, as $\mu^{(r)} \to \lambda_s$, $\mathbf{M}^{(r)}$ becomes increasingly near-singular and correspondingly ill-conditioned, and its reciprocal $\mathbf{N}^{(r)}$ grows accordingly, and therefore the equation (9.2) becomes increasingly difficult to solve accurately. This difficulty is intrinsic and puts a limit on the accuracy obtainable by this method.

Secondly, it is sometimes difficult to get the eigenvalue and eigenvector we choose, unless we begin rather close to the desired values. This is a familiar problem with iterative methods [compare the classic inverse power method, or Newton's method].

10. ACKNOWLEDGEMENT

I am indebted to the Los Alamos National Laboratory for organizing, and inviting me to participate in, a Workshop on Adaptive Monte Carlo Methods at their Center for Non-Linear Systems, in August 1996. At this workshop, I had the opportunity to hear of recent work and to present some of my own. More particularly, I am happy to have had the opportunity to discuss some of the material presented in §4–§9 of this present paper, especially with Dr T. E. Booth, Dr A. R. Forster, and Dr. T. T. Warnock of LANL, and Dr E. W. Larsen of the University

of Michigan at Ann Arbor. They will doubtless recognize my efforts to find satisfactory answers for their stimulating questions; any remaining shortcomings are all mine.

11. REFERENCES

W. F. AMES, 1992. *Numerical Methods for Partial Differential Equations*, Third Edition, Academic Press, New York (1992) 451 pp.

J.-P. AUBIN, 1972. *Approximation of Elliptic Boundary Value Problems.* Wiley-Interscience, New York, NY (1972) 360 pp.

O. AXELSSON, V. A. BARKER, 1984. *Finite Element Solution of Boundary Value Problems—Theory and Computation,* Academic Press, New York (1984) 432 pp.

K. BAGGERLY, 1996. Theory for exponential convergence, Part III: interfacing theory and practice (invited presentation at *LANL / CNLS Workshop on Adaptive Monte Carlo Methods,* Los Alamos National Laboratory, Los Alamos, NM, August 1996).

R. BEAUWENS, P. DE GROEN, 1992. Editors, *Iterative Methods in Linear Algebra* (Proceedings of the IMACS International Symposium, Brussels, Belgium, 1991), North-Holland/Elsevier, Amsterdam, The Netherlands (1992) 636 pp.

T. E. BOOTH, 1996. Adaptive Monte Carlo attempts on a continuous transport problem (invited presentation at *LANL / CNLS Workshop on Adaptive Monte Carlo Methods*, Los Alamos National Laboratory, Los Alamos, NM, August 1996).

N. P. BUSLENKO, D. I. GOLENKO, YU. A. SHREIDER, I. M. SOBOL', V. G. SRAGOVICH, 1962. *The Method of Statistical Trials—The Monte Carlo Method,* edited by YU. A. SHREIDER; Fizmatgiz, Moscow, USSR (1962) [in Russian]; Elsevier, Amsterdam, Netherlands (1964) 312 pp.; Pergamon Press, Oxford, England (1966) 390 pp.

L. L. CARTER, E. D. CASHWELL, 1975. *Particle Transport Simulation with the Monte Carlo Method,* Tech. Inf. Ctr., ERDA, Oak Ridge, TN (1975) 121 pp.

R. COURANT, K. O. FRIEDRICHS, H. LEWY, 1928. On the partial difference equations of mathematical physics, *Math. Ann.* 100 (1928) pp. 32–74 [in German].

R. R. COVEYOU, 1960. Serial correlation in the generation of pseudo-random numbers, *ACM Journal* 7 (1960) pp. 72–74.

D. COX, 1996. Theory for exponential convergence, Part II: continuous state spaces (invited presentation at *LANL/CNLS Workshop on Adaptive Monte Carlo Methods*, Los Alamos National Laboratory, Los Alamos, NM, August 1996).

J. H. CURTISS, 1949. Sampling methods applied to differential and difference equations, *Seminar on Sci. Comp., IBM Corp.,* New York, N Y (1949) pp. 87–109.

J. H. CURTISS, 1954. Monte Carlo methods for the iteration of linear operators, *J. Math. Phys.* 32 (1954) pp. 209–232.

R. E. CUTKOSKY, 1951. A Monte Carlo method for solving a class of integral equations, *J. Res. Nat. Bur. Stand.* 47 (1951) pp. 113–115.

G. DAHLQUIST, Å. BJÖRCK, 1974. *Numerical Methods,* Prentice-Hall, Englewood Cliffs, NJ (1974) 573 pp.

H. P. EDMUNDSON, 1953. Monte Carlo matrix inversion and recurrent events, *Math. Tab. Aids Comp.* 7 (1953) pp. 18–21.

S. M. ERMAKOV, 1975. *The Monte Carlo Method and Contiguous Questions,* Nauka, Moscow, USSR; First Edition (1971) 328 pp.; Second Edition (1975) 472 pp. [in Russian].

G. E. FORSYTHE, R. A. LEIBLER, 1950. Matrix inversion by a Monte Carlo method, *Math. Tab. Aids Comp.* 4 (1950) pp. 127–129.

G. E. FORSYTHE, C. B. MOLER, 1967. *Computer Solution of Linear Algebraic Systems.* Prentice-Hall, Inc., Englewood Cliffs, NJ (1967) 148 pp.

G. E. FORSYTHE, W. R. WASOW, 1960. *Computer Solution of Linear Algebraic Systems,* John Wiley & Sons, Inc., New York (1960) 444 pp.

P. W. G. GLYNN, 1996. On Markov processes and adaptive algorithms (invited presentation at *LANL/CNLS Workshop on Adaptive Monte Carlo Methods*, Los Alamos National Laboratory, Los Alamos, NM, August 1996).

J. H. HALTON, 1962. Sequential Monte Carlo, *Proc. Camb. Phil. Soc.* 58 (1962) pp. 57–78.

J. H. HALTON, 1965a. A general formulation of the Monte Carlo method and a 'strong law' for certain sequential schemes, *Brookhaven National Laboratory,* AMD 378/BNL 9220 (1965) 14 pp.

J. H. HALTON, 1965b. Least-squares Monte Carlo methods for solving linear systems of equations, *Brookhaven National Laboratory,* AMD 388/BNL 9678 (1965) 74 pp.

J. H. HALTON, 1966. On the strong convergence of linear averages, *Univ. Wisconsin, Madison,* MRC 719 (1966) 8 pp.

J. H. HALTON, 1967. Sequential Monte Carlo (Revised), *Univ. Wis., Madison,* MRC 816 (1967) 38 pp.

J. H. HALTON, 1970, A retrospective and prospective survey of the Monte Carlo method, *SIAM Review* 12 (1970) pp. 1–63.

J. H. HALTON, 1990. Monte Carlo methods for solving linear systems of equations, 18 pp. (invited presentation at *NSF–CBMS Res. Conf. Random Number Generation / Quasi-Monte-Carlo Methods,* Fairbanks, Alaska, August 1990).

J. H. HALTON, 1991a. The Monte Carlo solution of linear systems. *Univ. North Carol., Chapel Hill,* Working Paper (1991) 132 pp.; reprinted in J. H. Halton, *Readings on the Monte Carlo Method* (1992) 404 pp.

J. H. HALTON, 1991b. An introduction to the Monte Carlo solution of linear systems, 18 pp. (invited presentation at *IMACS International Symp. Iterative Methods in Linear Algebra*, Brussels, Belgium, April 1991).

J. H. HALTON, 1991c. Some new results on the Monte Carlo solution of linear systems, including sequential methods, 20 pp. (invited presentation
at *IMACS International Symp. Iterative Methods in Linear Algebra*, Brussels, Belgium, April 1991).

J. H. HALTON, 1994. Sequential Monte Carlo techniques for the solution of linear systems. *Journal of Scientific Computing* 9 (1994) pp. 213–257.

J. H. HALTON, 1996. Sequential Monte Carlo techniques for solving nonlinear systems, 35 pp. (invited presentation at *LANL / CNLS Workshop*
on Adaptive Monte Carlo Methods, Los Alamos National Laboratory, Los Alamos, NM, August 1996).

J. M. HAMMERSLEY, D. C. HANDSCOMB, 1964. *Monte Carlo Methods,* Methuen, London, England; John Wiley & Sons, New York (1964) 185 pp.

A. S. HOUSEHOLDER, 1964. *The Theory of Matrices in Numerical Analysis*, Blaisdell Publishing Co., New York, NY (1964) 257 pp.

E. ISAACSON, H. B. KELLER, 1966. *Analysis of Numerical Methods*, John Wiley & Sons, New York (1966) 541 pp.

M. H. KALOS, P. A. WHITLOCK, 1986. *Monte Carlo Methods, Volume I: Basics,* John Wiley & Sons, New York (1986) 186 pp.

J. P. C. KLEIJNEN, 1974. *Statistical Techniques in Simulation,* Marcel Dekker, New York, *Part I* (1974) 300 pp.

J. P. C. KLEIJNEN, 1975. *Statistical Techniques in Simulation,* Marcel Dekker, New York, *Part II* (1975) 503 pp.

E. W. LARSEN, 1996. An adaptive Monte Carlo method for global neutron transport calculations (invited presentation at *LANL / CNLS Workshop on Adaptive Monte Carlo Methods*, Los Alamos National Laboratory, Los Alamos, NM, August 1996).

A. W. MARSHALL, 1956. The use of multi-stage sampling schemes in Monte Carlo, *Univ. Fla., Gainesville, Symp. Monte Carlo Methods, 1954*, edited by H. A. MEYER; John Wiley & Sons, New York (1956) pp. 123–140.

G. W. MCKINNEY, R. KONG, 1996. Error reduction using adaptive Monte Carlo—the reduced source method (invited presentation at *LANL / CNLS Workshop on Adaptive Monte Carlo Methods*, Los Alamos National Laboratory, Los Alamos, NM, August 1996).

M. E. MULLER, 1956a. Some continuous Monte Carlo methods for the Dirichlet problem, *Ann. Math. Stats.* 27 (1956) pp. 569–589.

M. E. MULLER, 1956b. On discrete operators connected with the Dirichlet problem, *J. Math. Phys.* 35 (1956) pp. 89–113.

J. M. ORTEGA, 1988. *Introduction to Parallel and Vector Solution of Linear Systems,* Plenum Press, New York, NY (1988) 305 pp.

J. M. ORTEGA, W. G. POOLE, 1981. *An Introduction to Numerical Methods for Differential Equations,* Pitman Publishing, Inc., Marshfield, MA (1981) 329 pp.

J. M. ORTEGA, W. C. RHEINBOLDT, 1970. *Iterative Solution of Nonlinear Equations in Several Variables,* Academic Press, Inc., New York, NY (1970) 572 pp.

A. M. OSTROWSKI, 1966. *Solution of Equations and Systems of Equations,* Academic Press, Inc., New York, NY (1966) 338 pp.

E. S. PAGE, 1954. The Monte Carlo solution of some integral equations, *Proc. Camb. Phil. Soc.* 50 (1954) pp. 414–425.

R. PICARD, 1996. Theory for exponential convergence, Part I: discrete state problems and interpolated smoothing (invited presentation at *LANL / CNLS Workshop on Adaptive Monte Carlo Methods*, Los Alamos National Laboratory, Los Alamos, NM, August 1996).

S. PISSANETZKY, 1984. *Sparse Matrix Technology*, Academic Press, Inc., London, England (1984) 321 pp.

J. K. REID, 1971. On the method of conjugate gradients for the solution of large sparse systems of linear equations, *Large Sparse Sets of Linear Equations* (Proceedings of the Oxford Conference of the Institute
of Mathematics and Its Applications, April 1970), edited by J. K. REID; Academic Press, London, England (1971) pp. 231–254.

W. C. RHEINBOLDT, 1974. *Methods for Solving Systems of Nonlinear Equations,* SIAM, Philadelphia, PA (1974) 104 pp.

R. D. RICHTMYER, K. W. MORTON, 1967. *Difference Methods for Initial-Value Problems,* Interscience Publishers, Second Edition, New York (1967)
405 pp.

R. Y. RUBINSTEIN, 1981. *Simulation and the Monte Carlo Method,* John Wiley & Sons, New York (1981) 293 pp.

Y. SAAD, 1996. *Iterative Methods for Sparse Linear Systems*, PWS Publishing Co., Boston, MA (1996) 447 pp.

I. M. SOBOL', 1973. *Monte Carlo Computational Methods,* Nauka, Moscow, USSR, (1973) 312 pp. [in Russian].

J. SPANIER, 1996. General sequential sampling methods for Monte-Carlo and quasi-Monte-Carlo applications (invited presentation at *LANL/CNLS Workshop on Adaptive Monte Carlo Methods*, Los Alamos National Laboratory, Los Alamos, NM, August 1996).

J. SPANIER, E. M. GELBARD, 1969. *Monte Carlo Principles and Neutron Transport Problems,* Addison-Wesley, Reading, MA (1969) 248 pp.

G. W. STEWART, 1973. *Introduction to Matrix Computations,* Academic Press, New York (1973) 441 pp.

J. F. TRAUB, 1964. *Iterative Methods for the Solution of Equations,* Prentice-Hall, Inc., Englewood Cliffs, NJ (1964) 310 pp.

R. S. VARGA, 1962. *Matrix Iterative Analysis,* Prentice-Hall, Englewood Cliffs, NJ (1962) 322 pp.

A. WALD, J. WOLFOWITZ, 1950. Bayes solutions of sequential decision problems, *Ann. Math. Stats.* 21 (1950) pp. 82–89

W. WASOW, 1951a. Random walks and the eigenvalues of elliptic difference equations, *J. Res. Nat. Bur. Stand.* 46 (1951) pp. 65–73.

W. WASOW, 1951b. On the mean duration of random walks, *J. Res. Nat. Bur. Stand.* 46 (1951) pp. 462–471.

W. WASOW, 1951c. On the duration of random walks, *Ann. Math. Stats.* 22 (1951) pp. 199–216.

W. WASOW, 1952. A note on the inversion of matrices by random walks, *Math. Tab. Aids Comp.* 6 (1952) pp. 78–81.

J. H. WILKINSON, 1965. *The Algebraic Eigenvalue Problem,* Clarendon Press (Oxford University Press), Oxford, England (1965) 662 pp.

J. WOLFOWITZ, 1946. On sequential binomial estimation, *Ann. Math. Stats.*
17 (1946) pp. 489–493.

J. WOLFOWITZ, 1947. The efficiency of sequential estimates and Wald's equation for sequential processes, *Ann. Math. Stats.* 18 (1947) pp. 215–230.

S. J. YAKOWITZ, 1977. *Computational Probability and Simulation,* Addison-Wesley, Reading, MA (1977) 262 pp.

D. M. YOUNG, 1971. *Iterative Solution of Large Linear Systems,* Academic Press, Inc., New York, NY (1971) 570 pp.

D. ZWILLINGER, 1992. *Handbook of Differential Equations,* Academic Press, Inc., San Diego, CA, Second Edition (1992) 787 pp.