REVIEWS AND DESCRIPTIONS OF TABLES AND BOOKS

The numbers in brackets are assigned according to the American Mathematical Society classification scheme. The 2000 Mathematics Subject Classification can be found in print starting with the 1999 annual index of *Mathematical Reviews*. The classifications are also accessible from `www.ams.org/msc/`.

1[49K10, 49K15, 49M05, 49M37, 49N10, 49N25, 65F10, 65K05, 65K10, 90C05, 90C20, 90C25, 90C30]—*Control perspectives on numerical algorithms and matrix problems*, by Amit Bhaya and Eugenius Kaszkurewicz, SIAM, Philadelphia, PA, 2006, xxvi + 272 pp., softcover, US$97.00; SIAM member US$67.00, ISBN 0-89871-602-0

As the title of this book suggests, it studies numerical algorithms by viewing them as control systems with state feedback, which achieve convergence via descent on a certain Lyapunov/objective function. For example, the continuous-time Newton algorithm

$$\frac{dx}{dt} = -\nabla f(x)^{-1} f(x), \qquad x(0) = x_0$$

for finding a zero of a differentiable $f : \Re^n \to \Re^n$ may be viewed as feedback control on the dynamical system

$$\frac{dx}{dt} = u$$

with the control law $u(x) = -\nabla f(x)^{-1} f(x)$ and the Lyapunov function $V(x) = \|f(x)\|^2$ (Section 2.1). The discrete-time Newton algorithm

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k)^{-1} f(x_k), \quad k = 0, 1, \ldots,$$

with stepsize $\alpha_k > 0$ may be similarly viewed. For an optimizer like this reviewer, it is interesting to see some of the alternative choices for the direction inspired by control, such as

$$-\nabla f(x)^{-1} \mathrm{sign}(f(x)) \qquad \text{and} \qquad -\mathrm{sign}(\nabla f(x)^T f(x)).$$

However, these directions lack the fast (superlinear) convergence property of the Newton direction in a neighborhood of a solution x^* with $\nabla f(x^*)$ nonsingular. In the optimization literature, the discrete-time Newton algorithm is known as the Gauss-Newton algorithm for minimizing $\|f(x)\|^2$ and various techniques have been developed to ensure robustness and fast convergence. In Section 2.3, standard methods for solving linear systems (minimal residual, Krylov subspace, conjugate gradient) are similarly interpreted. Chapter 3 continues with the optimal control perspective in investigating related numerical algorithms for zero finding and unconstrained optimization. Chapter 4 studies gradient-based continuous-time algorithms, motivated by neural network training, for solving linear equations, convex programs, and linear programs. Additional topics in the numerical solution of ODEs and in matrix theory are presented in Chapter 5.

It is an intriguing premise to interpret numerical algorithms from a control perspective. Does it add new insights into the numerical algorithms? This remains

to be seen. The book should appeal to anyone interested in seeing some numerical algorithms being reinterpreted from a control perspective. However, if one is interested in practical algorithms for optimization and matrix computation, then the following books might be more useful: *Matrix Computation* (by Golub and Van Loan), *Numerical Linear Algebra* (by Trefethen and Bau), *Nonlinear Programming* (by Bertsekas), *Numerical Optimization* (by Nocedal and Wright), *Practical Methods of Optimization* (by Fletcher). For example, for unconstrained optimization, quasi-Newton algorithms are generally the most efficient and various strategies have been developed to adaptively choose the stepsize α_k to achieve robustness and fast convergence. For linear programming, dual simplex and primal-dual interior-point algorithms are currently the most efficient and can solve problems with millions of variables. For Support Vector Machines training, interior-point algorithms and block-coordinate minimization algorithms are currently the most efficient and can solve problems with 100,000 data points.

Paul Tseng
Department of Mathematics
University of Washington
Seattle, WA 98195