robust standard errors. But those with access to the appropriate matrix computations should probably compute the jack-knifed HC3. While the ease of doing this varies from package to package, it is surely easy enough to do so in `LIMDEP`,

Finally, it should be noted that White's approach to standard errors which are robust to heteroskedasticity succeeds because it does not assume that the analyst knows the nature of the heteroskedasticity. Such ignorance is clearly the most common situation. But there are times, such as with time-series–cross-section data, that the analyst may have some better insight about the nature of $\mathbf{\Omega}$. Such structure can then be incorporated into Equation 3. This is the basis for the "panel correct standard errors" I developed with Jonathan Katz (Beck and Katz 1995). There are also circumstances where knowledge about the form of the heteroskedasticity can be used to improve estimation through weighted least squares. Such an approach has proven extremely useful in the analysis of time-series, where it is often the case that heteroskedasticity follows a simple autoregressive form, leading to Engle's (1982) autoregressive conditional heteroskedasticity (ARCH) model and its generalizations. But in most cross-sectional studies it is hard to parameterize heteroskedasticity. In such cases the computation of robust standard errors at least lessens the likelihood of incorrect inference.

References

Beck, Nathaniel and Jonathan N. Katz. 1995. "What To Do (and Not To Do) with Times-Series Cross-Section Data." *American Political Science Review* 89:634-47.

Davidson, Russell and James MacKinnon. 1993. *Estimation and Inference in Econometrics.* New York: Oxford University Press.

Engle, Robert. 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50:987-1008.

Greene, William H. 1993. *Econometric Analysis.* Second edition. New York: Macmillan.

Greene, William H. 1995. *LIMDEP: Version 7.0 Users Manual.* Bellport, N.Y.: Econometric Software, Inc.

Leamer, Edward. 1978. *Specification Searches: Ad Hoc Inference with Non-experimental Data.* New York: Wiley.

MacKinnon, James and Halbert White. 1984. "Some Heteroskedasticity-consistent Covariance Matrix Estimators with Improved Finite Sample Properties." *Journal of Econometrics* 29:305=325.

White, Halbert. 1980. "A Heteroscedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity." *Econometrica* 48:817-838.

Nathaniel Beck is Professor of Political Science at UCSD. He is currently working on time-series–cross-section models, the application of generalized additive models and the analysis of discrete duration data. These techniques are used to study comparative political economy and international relations.

A Review of Discrete Optimization Algorithms

Andrew D. Martin and Kevin M. Quinn
Washington University
admartin@wugate.wustl.edu and kquinn@wugate.wustl.edu

Introduction

Political scientists are often faced with optimization problems involving discrete parameter spaces, multi-modal functions, functions which are not well defined, noisy functions, and even non-differentiable functions. These problems can arise in maximum likelihood estimation (MLE), forecasting, dynamic modeling, and some types of game theoretic models. While these problems were daunting 10 years ago, recent advances in computational optimization combined with falling computer prices have brought us much closer to reasonable solutions. In these applications, we encounter optimization problems of the following general form: minimize $c(x_i)$ s.t. $x_i \in \mathbf{X}$ where $c(\cdot)$ is the objective function and $\mathbf{X}$ is the solution space[1]. The objective function $c(\cdot)$ maps the members of the solution space onto the real number line. This article serves as a review of three widely used discrete optimization algorithms that are well suited to dealing with the problems above and provides suggestions of the types of problems each method is well suited. The three techniques we review are: genetic algorithms, tabu search, and simulated annealing. These algorithms are able to provide solutions to optimization problems in which calculus based optimization is infeasible or impossible. While the focus of this review is on discrete optimization, it should be noted that two of the algorithms

[1] Note that minimizing $c(\cdot)$ is equivalent to maximizing $-c(\cdot)$.

(genetic algorithms and simulated annealing) discussed below have real-coded counter-parts.

Genetic Algorithms

Genetic algorithms (GAs) were created by Holland (1975, 1992) to study the mathematical underpinnings of adaptive behavior. While GAs were designed to simulate natural adaptive behavior, they have also proven to be very powerful optimization procedures. This twofold ability of GAs to both simulate adaptive behavior and to solve extremely difficult optimization problems has proven quite useful to several social scientists. Examples of social scientific work employing GAs include searches for optimal strategies in complicated formal models (Axelrod 1987; Miller 1989; Andreoni and Miller 1990; and Kollman, Miller, and Page 1992), and the estimation of LISREL models (Mebane et al. 1995). GAs have also been used to find solutions to systems of nonlinear equations (Shaefer, 1985).

GAs are implemented in roughly the following manner. First, choose a coding scheme which maps each element of the search space onto a unique bit vector [2]. Then randomly select a small set, $X_t \subset \mathbf{X}$, of m potential solutions to the optimization problem. Evaluate the objective function $c(\cdot)$ at each potential solution $x_i \in X_t$. The difference between $c(x_i)$ and some measure of the expected value of $c(x_i)$ (for example, $\frac{1}{m}\sum_{j=3D1}^{m} c(x_j)$) provides a measure of fitness for x_i. Then form a new set of potential solutions, X_{t+1}, by applying a series of genetic operators to X_t. Reproduction generally occurs first. This operator selects potential solutions from X_t (with replacement) on the basis of fitness (relatively more fit solutions are more likely to be selected). The crossover operator is then applied. The crossover operator randomly takes vectors of potential solutions, breaks them apart, and recombines them. For example, the crossover operator could transform {11111} and {00000} into {11000} and {00111}. Finally, the mutation operator flips elements of the solution vectors to the opposite value with some low probability. Denote the new set of solutions formed from X_t as $X_{t+1} \subset \mathbf{X}$.

Iterating this procedure a number of times yields a very powerful optimization algorithm. As Holland and Miller (1991) note, this three-part process of reproduction, crossover, and mutation may seem to be nothing more than a random search algorithm which retains the best potential solutions. As they further argue, this is in fact not the case. In order to understand why, think of each bit of a solution vector as an arm of an n-armed bandit. The problem then is to allocate trials to each of the n arms in a manner which will yield the highest cumulative payoff. Holland (1975, 1992) has shown that GAs allocate trials to building blocks in a manner which very closely corresponds to the optimal solution of an n-armed bandit problem.

Goldberg (1989) provides a non-technical introduction to GAs. For advanced discussions of additional genetic operators as well as refinements and variations of the three basic genetic operators we urge the reader to see the edited volumes by Grefenstette (1987), Rawlins (1991), and Whitley (1993).

Tabu Search

Tabu search, first proposed by Glover (1977), is a meta-heuristic used to solve both combinatorial and discrete optimization problems. For the purpose of discrete optimization problems, the heuristic used in the tabu search algorithm is a local improvement scheme, beginning with a good feasible solution. Local search starts from an initial solution $x_i \in \mathbf{X}$ and searches to find an improving solution $x_{i+1} \in \mathbf{X}$. In other words, the search attempts to find an x_{i+1} such that $c(x_{i+1}) < c(x_i)$.

Consider the case when optimizing over a discrete space $\mathbf{X}$ with respect to an *a priori* objective function $c(\cdot)$. Define $\mathbf{X}$ as the solution space which contains all of the possible solutions $x \in \mathbf{X}$. For each x_i the practitioner defines a set $N(x_i) \subset \mathbf{X}$ which denotes the neighborhood of x_i. On each iteration of the search, the objective function value is evaluated for all $x \in N(x_i)$. The entire neighborhood is searched and an improving move is chosen by selecting the move x_i which most improves the value of $c(x_i)$. That move x_i is chosen and is labeled x_{i+1}. The search moves to the next iteration, by looking at the neighborhood of the accepted move, $N(x_{i+1})$. The problem with simple local search is that traps of local optimality cannot be escaped. In a discrete space, local optimality is defined in terms of an *a priori* neighborhood structure as opposed to an $\varepsilon-$neighborhood in the continuous case. Although unimodal functions are easily optimized, multi-modal functions make local search an impractical optimization technique. To remedy this problem, simple local search is modified to accept some non-improving moves in an attempt to escape traps of local optimality.

The first of these meta-heuristics based on local search is called tabu search. Tabu search uses a memory structure (called the tabu list) that restricts the possible members of the neighborhood to which a search can progress. Thus, once a local optima is encountered, the search will not be able to revisit that area of the solution space. The tabu list must be small enough to allow the search to carefully scrutinize certain parts of the solution

[2] Real number encoded GAs are also possible. For discussions of the implementation and performance of real-coded GAs, we refer the reader to Eshelman and Schaffer (1993), Wright (1991), and Antonisse (1989). While we only discuss binary-coded GAs here for reasons of simplicity and space, it should be noted that the use of real-coded GAs has grown rapidly in the pas t few years.

space, yet large enough to prevent a return to a previously generated solution. The tabu search meta-heuristic also uses an aspiration criterion which defines a condition under which the tabu status of a certain move can be overridden. Short term memory functions are employed to intensify and diversify the search. Tabu search is allowed to run for a maximum number of iterations that is computationally practical. A comprehensive description of tabu search can be found in Glover and Laguana (1993).

When implementing tabu search, the practitioner must define the neighborhood structure with respect to the solution space, select the type of tabu list to be employed, and determine the aspiration criterion to be used. Practitioners also traditionally choose to employ multi-start techniques, where tabu search is re-started numerous times from different members of th e solution space. Throughout the operations research literature, there are many examples of successful implementations of tabu search as well as discussions of effective tabu structures and aspiration criteria (Cvijovic and Klinowski 1995; Glover and Laguana 1993; and Glover 1990).

Simulated Annealing

Another meta-heuristic which relies on local search is called simulated annealing. Simulated annealing was first introduced by Kirkpatrick et. al. (1983) and Cerny (1985) and has roots in the work of Metropolis et al. (1953). Simulated annealing is analogous to the annealing process in physical chemistry, when liquid metals are heated and then left to cool into a steady, organized state. Numerous successful applications of simulated annealing can be found in Collins, et. al. (1988). The simulated annealing algorithm can be described in terms of a Markov chain. The solution space $\mathbf{X}$ consists of the feasible solutions that satisfy all the constraints $x \in \mathbf{X}$. An objective function $c(\cdot)$ is defined on $\mathbf{X}$. From each state x_i a transition is a search action that combines the selection of a state $x_j \in N(x_i)$ with the decision of whether to move to x_j state or not. The neighborhood $N(x_i) \subset \mathbf{X}$ of state x_i is defined as the set of states that can be reached from state x_i in exactly one step. Thus, if the transition probability $p_{x_i x_j} > 0$, then $x_i \in N(x_j)$. Furthermore, the selection is reversible; if $x_j \in N(x_i)$ then $x_i \in N(x_j)$. In simulated annealing, each member of the neighborhood is randomly selected, and the algorithm then determines whether to move on to the next state. The decision to move to the next state depends on the values of $c(x_i)$ and $c(x_{i+1})$. The decision allows the acceptance of some non-improving moves, thus escaping traps of local optimality. The algorithm provides a chance for the search to escape from a local optimum based on an acceptance probability, which is defined as:

$$\Pr(accept\ x_{i+1}) = min\left\{1, \frac{\exp[c(x_{i+1}) - c(x_i)]}{t}\right\} \quad (1)$$

where t is the temperature control parameter. This temperature control parameter is decreased as the search progresses, thus allowing the search to settle down into a local optimum.

It can be shown that convergence to the global optimum is guaranteed if the temperature control parameter approaches 0 and an infinite number of transitions is made. However, since this convergence is quite impractical, a finite-time implementation of the simulated annealing algorithm is often used to approach the optimal solution within a reasonable amount of computation time. When implementing simulated annealing, a practitioner must define the neighborhood structure with respect to the solution space and develop a cooling schedule with which to decrement the temperature parameter t. A survey of successful cooling schedules can be found in Hajek (1988) and Collins et. al. (1988).

Simulated annealing is guaranteed to converge to the global optimum of functions defined over both discrete and continuous spaces as the cooling parameter t goes to zero. Thus, it is particularly appropriate for estimation of econometric models.[3]

Conclusion

In sum, the operations research literature provides numerous computational techniques that political scientists can implement to conquer previously intractable problems. These techniques can be applied to optimization problems encountered in the estimation of econometric models, forecasting, dynamic modeling, and some types of game theoretic models. For a comprehensive description, evaluation, and comparison of many discrete optimization techniques, we refer the reader to Ackley (1987) who empirically assesses the success of each algorithm.

References

Ackley, David H. 1987. "An Empirical Study of Bit Vector Optimization." in Lawrence D. Davis (ed.). *Genetic Algorithms and Simulated Annealing.* London: Pittman.

Andreoni, James, and John H. Miller. 1990. "Auctions Wit h Artificial Adaptive Agents." University of Wisconsin Department of Economics, mimeo.

[3]It is interesting to note that the underlying logic behind the method of simulated annealing can also be used to simulate multivariate distributions. See Chib and Greenberg (1995) for a more detailed discussion of these issues.

Antonisse, H. J. 1989. "A New Interpretation of Schema Notation that Overturns the Binary Encoding Constraint." *Proceedings of the Fourth International Conference on Genetic Algorithms.* San Mateo, CA: Morgan Kaufmann.

Axelrod, Robert. 1987. "The Evolution of Strategies in the Iterated Prisoner's Dilemma." in Lawrence D. Davis (ed.). *Genetic Algorithms and Simulated Annealing.* London: Pittman.

Cerny, V. 1985. "Thermodynamical Approach to the Traveling Salesman Problem: An Efficient Simulation Program." *Journal of Optimization Theory and Applications.* 45: 41-51.

Chib, Siddhartha and Edward Greenberg. 1995. "Understanding the Metropolis-Hastings Algorithm." *The American Statistician.* 49: 327-335.

Collins, N.E., R.W. Eglese and B. L. Golden. 1988. "Simulated Annealing - An Annotated Bibliography." *American Journal of Mathematical and Management Sciences.* 8: 209-307.

Cvijovic, Djurdje, and Jacek Klinowski. 1995. "Taboo Search: An Approach to the Multiple Minima Problem." *Science.* 267: 664-666.

Eshelman, L. J., and J. D. Schaffer. 1993. "Real-Coded Genetic Algorithms and Interval-Schemata." in L. Darrell Whitley (ed.). *Foundations of Genetic Algorithms 2.* San Mateo, CA: Morgan Kaufmann .

Glover, Fred. 1977. "Heuristics for Integer Programming Using Surrogate Constraints." *Decision Sciences.* 8: 156-166.

Glover, Fred. 1990. "Tabu Search: A Tutorial." *Interfaces.* 20: 74-94.

Glover, Fred and Manual Laguana. 1993. "Tabu Search." in Colin R. Reeves, (ed.) *Modern Heuristic Techniques for Combinatorial Problems.* Oxford: Blackwell Scientific Publications. pp. 70-150.

Goldberg, David E. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning.* Reading, MA: Addison Wesley.

Grefenstette, John J. (ed.) 1987. *Genetic Algorithms and Their Applications: Proceedings of the Second International Conference on Genetic Algorithms.* Hillsdale NJ,: Lawrence Erlbaum Associates.

Hajek, B. 1988. "Cooling Schedules for Optimal Annealing." *Mathematics of Operations Research.* 13: 311-329.

Holland, John H. [1975] 1992. *Adaptation in Natural and Artificial Systems.* Cambridge, MA: MIT Press.

Holland, John H., and John H. Miller. 1991. "Artificial Adaptive Agents in Economic Theory." *American Economic Review.* 81: 365-370..

Kirpatrick, S., C.D. Gelatt and M.P. Vecchi. 1983. "Optimization by Simulated Annealing." *Science.* 220: 671-680.

Kollman, Ken, John H. Miller, and Scott Page. 1992. "Adaptive Parties in Spatial Elections." *American Political Science Review.* 86: 929-937.

Mebane, Walter R., Jasjeet S. Sekhon, and Martin T. Wells. 1995. "The Robustness of Normal-theory LISREL Models: Tests Using a New Optimizer, the Bootstrap, and Sampling Experiments, with Applications". Paper presented at the 1995 Political Methodology Summer Conference.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. "Equations of State Calculations by Fast Computing Machines." *Journal of Chemical Physics.* 21: 1087-1092.

Miller, John H. 1989. "The Co-Evolution of Automata in the Repeated Prisoners Dilemma." Sante Fe Institute Economics Research Program Working Paper No. 89-003.

Rawlins, G. J. E. (ed.) 1991. *Foundations of Genetic Algorithms.* San Mateo, CA: Morgan Kaufmann.

Shaefer, C. G. 1985. "Comparison of Methods for Solving Nonlinear Equations." Research Memo RIS-24r. Cambridge, MA: Rowland Institute for Science.

Whitley, Darrel L. (ed.). 1993. *Foundations of Genetic Algorithms 2.* San Mateo CA: Morgan Kaufmann.

Wright, A. 1991. "Genetic Algorithms for Real Parameter Optimization." in *Foundations of Genetic Algorithms.* G. J. E. Rawlings, ed. San Mateo, CA: Morgan Kaufmann.

Andrew D. Martin received an A.B. in mathematics and government from The College of William and Mary. He is currently pursuing a Ph.D. in political science at Washington University in St. Louis. His academic interests include American politics, formal theory, and political methodology.

Kevin M. Quinn received an B.A. in political science from The Johns Hopkins University and an A.M. in political science from Washington University in St. Louis, where he is presently a doctoral candidate. His

academic interests include comparative politics, formal theory, and political methodology.

Polmeth — You've Come a Long Way Baby

R. Michael Alvarez
California Institute of Technology
rma@crunch.caltech.edu

In many ways it is difficult to remember life without a direct pipeline into the Internet and the World Wide Web. It was just within the past few years that most people in political science have migrated full-force onto the Internet and the Web, and it is becoming quite clear now that both are shaping the way we engage in research and how we interact professionally, in quite profound ways.

One of the developments of interest to political methodologists has been the rapid evolution of our Political Methodology World Wide Web server and our "polmeth" discussion group, both maintained by Jonathan Nagler at the University of California, Riverside, through generous support by the National Science Foundation and UC Riverside. The purpose of this article is to take a brief look at the progress of *Polmeth* in the past year. I want to present some statistics on the usage of *Polmeth* which clearly document the dramatic and rapid effect which *Polmeth* has had on political methodology in the past year, and then present a few ideas for future development of our professional and research connections to the Web.

A Brief History

Polmeth began without much of a bang in the spring of 1994. A number of people began an email discussion that spring focused on both the desireability and the functionality of providing a centralized place where people could deposit the papers which they were to present at upcoming political methodology summer conferences, and at other national meetings. A number of important issues were raised in these discussions:

- Where would the paper repository be located, and how could it be maintained?
- How could we encourage (or worse yet, coerce) our colleagues into using this internet service instead of making endless copies of papers, hauling them on airplanes bound for their next meeting, and passing them out at the meeting instead of distributing them beforehand?
- What formats would we use? How could people easily distribute machine-readable versions of their papers without running the risk that the contents could be easily altered?

To examine the practical issues behind the development of a true Internet paper distribution system, we began two simple experiments. We convinced the section leadership that this practice would help advance methodology intellectually — and that they should thereby encourage people who were presenting their papers at the 1994 Summer Methodology Conference to provide machine-readable versions of their papers to the participants of the meeting, before the meeting. We set up an anonymous ftp directory at Caltech where paper presenters could place machine-readable versions of their papers, and an email reflector where they could send an email which would be "bounced" to every meeting participant.

To resolve the practical issues of paper format, we asked people to provide, at bare minimum, a version of their paper which could be printed on any HP laser jet. We also asked people to provide a Postscript version of their paper, if possible. Our thinking was that virtually everyone we knew had either an HP laser jet, or an Apple-style laserwriter. Thus, these two formats should cover the bases.

The experiment was a great success. Almost all of the papers presented at the 1994 Summer Methodology Conference were uploaded to our anonymous ftp server before the meeting; after the meeting another paper or two were added (they still are available on *Polmeth*!). As each paper was uploaded, we verified the integrity of the upload by checking that we ourselves could print the paper; if the upload was successful, we notified every meeting participant of the availability of the paper. The only problems we encountered in our experiment were "persuading" our colleagues that it was in their best interest to "post" their paper before the meeting, and some difficulties associated with the improper use of non-binary transmissions and retrievals. At the 1994 Conference, we had an open discussion of this experiment, and the meeting participants were virtually unified in their recommendation that we attempt to expand this service.

The Development of *Polmeth*

The rapid evolution of the Web in late 1994 and early 1995 facilitated this task. The members of our informal discussion group agreed that the development of a Web server for the Political Methodology community was the right direction to take. With the support of the NSF, Nagler was able to set up *Polmeth* on a high-speed Hewlett-Packard workstation in April 1995; after that, *Polmeth* was in business!

The early offerings on *Polmeth* were scarce. There were a series of links to other "interesting" sites, information about the 1995 Summer Methodology Conference, and