Supplement to "Numerically stable and accurate stochastic simulation approaches for solving dynamic economic models": Appendices

(*Quantitative Economics*, Vol. 2, No. 2, July 2011, 173–210)

Kenneth L. Judd
Hoover Institution, Stanford University and NBER

Lilia Maliar
University of Alicante and Hoover Institution, Stanford University

Serguei Maliar
University of Alicante and Hoover Institution, Stanford University

Appendix A: Nonlinear regression model and nonlinear approximation methods

In this section, we extend the approximation approaches that we developed in Sections 4.2 and 4.3 to the case of the nonlinear regression model

$$y = \Psi(k, a; b) + \varepsilon, \tag{A.1}$$

where $b \in \mathbb{R}^{n+1}$, $k \equiv (k_0, \dots, k_{T-1}) \in \mathbb{R}^T$, $a \equiv (a_0, \dots, a_{T-1}) \in \mathbb{R}^T$, and $\Psi(k, a; \beta) \equiv (\Psi(k_0, a_0; \beta), \dots, \Psi(k_{T-1}, a_{T-1}; \beta))^\top \in \mathbb{R}^T$.[1] We first consider a nonlinear LS (NLLS) problem and then formulate the corresponding LAD problem.

The NLLS problem is

$$\min_b \|y - \Psi(k, a; b)\|_2^2 = \min_b [y - \Psi(k, a; b)]^\top [y - \Psi(k, a; b)]. \tag{A.2}$$

The typical NLLS estimation method linearizes (A.2) around a given initial guess b by using a first-order Taylor expansion of $\Psi(k, a; b)$ and makes a step Δb toward a solution

$$\widehat{b} \simeq b + \Delta b. \tag{A.3}$$

Using the linearity of the differential operator, we can derive an explicit expression for the step Δb. This step is given by a solution to the system of normal equations

$$\mathcal{J}^\top \mathcal{J} \Delta b = \mathcal{J}^\top \Delta y, \tag{A.4}$$

Kenneth L. Judd: kennethjudd@mac.com
Lilia Maliar: maliarl@stanford.edu
Serguei Maliar: maliars@stanford.edu

[1] The regression model with the exponentiated polynomial $\Psi(k_t, a_t; b) = \exp(b_0 + b_1 \ln k_t + b_2 \ln a_t + \cdots)$, used in Marcet's (1988) simulation-based PEA, is a particular case of (A.1).

Copyright © 2011 Kenneth L. Judd, Lilia Maliar, and Serguei Maliar. Licensed under the Creative Commons Attribution-NonCommercial License 3.0. Available at http://www.qeconomics.org.
DOI: 10.3982/QE14

where $\mathcal{J}$ is a Jacobian matrix of Ψ,

$$\mathcal{J} \equiv \begin{pmatrix} \frac{\partial \Psi(k_0, a_0; b)}{\partial b_0} & \cdots & \frac{\partial \Psi(k_0, a_0; b)}{\partial b_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial \Psi(k_{T-1}, a_{T-1}; b)}{\partial b_0} & \cdots & \frac{\partial \Psi(k_{T-1}, a_{T-1}; b)}{\partial b_n} \end{pmatrix},$$

and

$$\Delta y \equiv (y_0 - \Psi(k_0, a_0; b), \ldots, y_{T-1} - \Psi(k_{T-1}, a_{T-1}; b))^\top.$$

Typically, the NLLS estimation method does not give an accurate solution $\widehat{b}$ in a single step Δb, and must instead iterate on the step (A.3) until convergence.[2]

A direct way to compute the step Δb from (A.4) is to invert the matrix $\mathcal{J}^\top \mathcal{J}$, which yields the well known *Gauss–Newton method*

$$\Delta b = (\mathcal{J}^\top \mathcal{J})^{-1} \mathcal{J}^\top \Delta y. \tag{A.5}$$

This formula (A.5) has a striking resemblance to the OLS formula $\widehat{b} = (X^\top X)^{-1} X^\top y$, namely, X, y, and b in the OLS formula are replaced in (A.5) by $\mathcal{J}$, Δy, and Δb, respectively. If $\mathcal{J}^\top \mathcal{J}$ is ill-conditioned, as is often the case in applications, the Gauss–Newton method experiences the same difficulties in computing $(\mathcal{J}^\top \mathcal{J})^{-1}$ and Δb that the OLS method does in computing $(X^\top X)^{-1}$ and b.

To deal with the ill-conditioned matrix $\mathcal{J}^\top \mathcal{J}$ in the Gauss–Newton method (A.5), we can employ the LS approaches similar to those developed for the linear regression model in Sections 4.2.1 and 4.2.2 of the paper. Specifically, we can compute an inverse of the ill-conditioned matrix $\mathcal{J}^\top \mathcal{J}$ by using LS methods based on SVD or QR factorization of $\mathcal{J}$. We can also use the Tikhonov type of regularization, which leads to the *Levenberg–Marquardt* method

$$\Delta b(\eta) = (\mathcal{J}^\top \mathcal{J} + \eta I_{n+1})^{-1} \mathcal{J}^\top \Delta y, \tag{A.6}$$

where $\eta \geq 0$ is a regularization parameter.[3]

Furthermore, we can replace the ill-conditioned NLLS problem (A.2) with a nonlinear LAD (NLLAD) problem

$$\min_b \|y - \Psi(k, a; b)\|_1 = \min_b 1_T^\top |y - \Psi(k, a; b)|. \tag{A.7}$$

As in the NLLS case, we can proceed by linearizing the nonlinear problem (A.7) around a given initial guess b. The linearized version of the NLLAD problem (A.7) is

$$\min_{\Delta b} 1_T^\top |\Delta y - \mathcal{J} \Delta b|. \tag{A.8}$$

The problem (A.8) can be formulated as a linear-programming problem: specifically, we can set up the primal and dual problems, as well as regularized primal and dual problems, analogous to those considered in Sections 4.3.1 and 4.3.2 of the paper.

[2]Instead of the first-order Taylor expansion of $\Psi(k, \theta; b)$, we can consider a second-order Taylor expansion, which leads to Newton's class of nonlinear optimization methods in which the step Δb depends on a Hessian matrix; see Judd (1992, pp. 103–117) for a review.

[3]This method was proposed independently by Levenberg (1944) and Marquardt (1963).

Example. Let us formulate a regularized primal problem for (A.8) that is parallel to (34)–(37) in the paper. Fix some initial φ^+ and φ^- (which determine initial $b(\eta) = \varphi^+ - \varphi^-$), and solve for $\Delta\varphi^+$ and $\Delta\varphi^-$ from the linear-programming problem

$$\min_{v^+, v^-, \Delta\varphi^+, \Delta\varphi^-} 1_T^\top v^+ + 1_T^\top v^- + \eta 1_n^\top \Delta\varphi^+ + \eta 1_n^\top \Delta\varphi^- \tag{A.9}$$

$$\text{s.t.} \quad v^+ - v^- + J\Delta\varphi^+ - J\Delta\varphi^+ = \Delta y, \tag{A.10}$$

$$v^+ \geq 0, \qquad v^- \geq 0, \tag{A.11}$$

$$\Delta\varphi^+ \geq 0, \qquad \Delta\varphi^- \geq 0. \tag{A.12}$$

Compute $\widehat{\varphi}^+ \simeq \varphi^+ + \Delta\varphi^+$ and $\widehat{\varphi}^- \simeq \varphi^- + \Delta\varphi^-$, and restore the regularized NLLAD estimator $\widehat{b}(\eta) \simeq (\varphi^+ + \Delta\varphi^+) - (\varphi^- + \Delta\varphi^-)$. As in the case of NLLS methods, we do not typically obtain an accurate solution $\widehat{b}$ in a single step, but must instead solve the problem (A.9)–(A.12) iteratively until convergence.

To set up a regularized dual problem for (A.8), which is analogous to (38)–(41) in the paper, we must replace X and y with $\mathcal{J}$ and Δy, respectively.

Notice that the NLLS and NLLAD regularization methods described in this section penalize all coefficients equally, including an intercept. Prior to applying these methods, we need to appropriately normalize the explanatory variables and to set the penalty on the intercept to zero.

Appendix B: Multidimensional deterministic integration methods

In this section, we describe deterministic integration methods suitable for evaluating multidimesional integrals of the form $\int_{\mathbb{R}^N} G(\epsilon) w(\epsilon)\, d\epsilon$, where $\epsilon \equiv (\epsilon^1, \ldots, \epsilon^N)^\top \in \mathbb{R}^N$ follows a multivariate Normal distribution $\epsilon \sim \mathcal{N}(\mu, \Sigma)$, where $\mu \equiv (\mu^1, \ldots, \mu^N)^\top \in \mathbb{R}^N$ is a vector of means and $\Sigma \in \mathbb{R}^{N\times N}$ is a variance–covariance matrix, and $w(\epsilon)$ is a density function of the multivariate Normal distribution,

$$w(\epsilon) = (2\pi)^{-N/2} \det(\Sigma)^{-1/2} \exp\left[-\frac{1}{2}(\epsilon - \mu)^\top \Sigma^{-1} (\epsilon - \mu)\right], \tag{B.1}$$

with $\det(\Sigma)$ denoting the determinant of Σ.[4]

B.1 *Cholesky decomposition*

The existing deterministic integration formulas are constructed under the assumption of uncorrelated random variables with zero mean and unit variance. If the random variables $\epsilon^1, \ldots, \epsilon^N$ are correlated, we must rewrite the integral in terms of uncorrelated variables prior to numerical integration. Given that Σ is symmetric and positive-definite, it has a Cholesky decomposition, $\Sigma = \Omega\Omega^\top$, where Ω is a lower triangular

[4] Such integration methods are used in Step 2 of GSSA to compute conditional expectation of the form $E_t\{G_t(\epsilon_{t+1})\} = \int_{\mathbb{R}^N} G_t(\epsilon) w(\epsilon)\, d\epsilon$ in each simulated point t, in particular, for the representative–agent model (2)–(4), $G_t(\epsilon_{t+1})$ is the integrand in (7).

matrix with strictly positive diagonal entries. The Cholesky decomposition of Σ allows us to transform correlated variables ϵ into uncorrelated ν with the linear change of variables

$$\nu = \frac{\Omega^{-1}(\epsilon - \mu)}{\sqrt{2}}. \tag{B.2}$$

Note that $d\epsilon = (\sqrt{2})^N \det(\Omega)\, d\nu$. Using (B.2), and taking into account that $\Sigma^{-1} = (\Omega^{-1})^\top \Omega^{-1}$ and that $\det(\Sigma) = [\det(\Omega)]^2$, we obtain

$$\int_{\mathbb{R}^N} G(\epsilon) w(\epsilon)\, d\epsilon = \pi^{-N/2} \int_{\mathbb{R}^N} G(\sqrt{2}\Omega\nu + \mu) \exp(-\nu^\top \nu)\, d\nu. \tag{B.3}$$

Deterministic integration methods approximate the integral (B.3) by a weighted sum of the integrand G evaluated in a finite set of nodes

$$\int_{\mathbb{R}^N} G(\epsilon) w(\epsilon)\, d\epsilon \approx \pi^{-N/2} \sum_{j=1}^{J} \omega_j G(\sqrt{2}\Omega\nu_j + \mu), \tag{B.4}$$

where $\{\nu_j\}_{j=1,\ldots,J}$ and $\{\omega_j\}_{j=1,\ldots,J}$ are integration nodes and integration weights, respectively. In the remaining section, we assume $\mu = 0_N$, where 0_N is an $N \times 1$ vector whose entries are equal to 0.

B.2 *Gauss–Hermite quadrature*

In a one-dimensional integration case, $N = 1$, the integral (B.4) can be computed using the Gauss–Hermite quadrature method. To be specific, we have

$$\int_{\mathbb{R}} G(\epsilon) w(\epsilon)\, d\epsilon \approx \pi^{-1/2} \sum_{j=1}^{J} \omega_j G(\sqrt{2}\Omega\nu_j), \tag{B.5}$$

where $\{\nu_j\}_{j=1,\ldots,J}$ and $\{\omega_j\}_{j=1,\ldots,J}$ can be found using a table of Gauss–Hermite quadrature nodes and weights (see, e.g., Judd (1998, p. 262)).

We can extend the one-dimensional Gauss–Hermite quadrature rule to the multidimensional case by way of a tensor-product rule

$$\begin{aligned} &\int_{\mathbb{R}^N} G(\epsilon) w(\epsilon)\, d\epsilon \\ &\quad \approx \pi^{-N/2} \sum_{j_1=1}^{J_1} \cdots \sum_{j_N=1}^{J_N} \omega_{j_1}^1 \cdots \omega_{j_N}^N \cdot G\big(\sqrt{2}\Omega \cdot (\nu_{j_1}^1, \ldots, \nu_{j_N}^N)^\top\big), \end{aligned} \tag{B.6}$$

where $\{\omega_{j_h}^h\}_{j_h=1,\ldots,J_h}$ and $\{\nu_{j_h}^h\}_{j_h=1,\ldots,J_h}$ are, respectively, weights and nodes in a dimension h derived from the one-dimensional Gauss–Hermite quadrature rule (note that, in general, the number of nodes in one dimension J_h can differ across dimensions). The total number of nodes is given by the product $J_1 J_2 \cdots J_N$. Assuming that $J_h = J$ for all

dimensions, the total number of nodes J^N grows exponentially with the dimensionality N.

B.3 *Monomial rules*

Monomial integration rules are nonproduct: they construct a relatively small set of nodes distributed in some way within a multidimensional hypercube. The computational expense of monomial rules grows only polynomially with the dimensionality of the problem, which makes them feasible for problems with large dimensionality.

We describe two monomial formulas for approximating the multidimensional integral (B.3). Monomial formulas are provided for the case of uncorrelated variables, for example, in Stroud (1971, pp. 315–329) and Judd (1998, p. 275). Here, we adapt them to the case of correlated random variables using the change of variables (B.2).

The first formula, denoted by $M1$, has $2N$ nodes,

$$\int_{\mathbb{R}^N} G(\epsilon)w(\epsilon)\,d\epsilon \approx \frac{1}{2N}\sum_{h=1}^{N} G(\pm R\iota^h), \tag{B.7}$$

where $R \equiv \sqrt{N}\Omega$, and $\iota^h \in \mathbb{R}^N$ is a vector whose hth element is equal to 1 and the remaining elements are equal to 0, that is, $\iota^h \equiv (0,\ldots,1,\ldots,0)^\top$.

The second formula, denoted by $M2$, has $2N^2+1$ nodes,

$$\begin{aligned}\int_{\mathbb{R}^N} G(\epsilon)w(\epsilon)\,d\epsilon \approx{}& \frac{2}{2+N}G(0,\ldots,0)\\ &+\frac{4-N}{2(2+N)^2}\sum_{h=1}^{N}[G(R\iota^h)+G(-R\iota^h)]\\ &+\frac{1}{(N+2)^2}\sum_{h=1}^{N-1}\sum_{s=h+1}^{N} G(\pm D\iota^h \pm D\iota^s),\end{aligned} \tag{B.8}$$

where $R \equiv \sqrt{2+N}\Omega$ and $D \equiv \sqrt{\frac{2+N}{2}}\Omega$.

B.4 *An example of integration formulas for $N=2$*

In this section, we illustrate the integration formulas described above using a two-dimensional example, $N=2$. We assume that the variables ϵ^1 and ϵ^2 are uncorrelated, have zero mean, and have unit variance. The integral (B.3) is then given by

$$E\{G(\epsilon)\} = \frac{1}{\pi}\int_{\mathbb{R}^2} G(\sqrt{2}\nu^1,\sqrt{2}\nu^2)\exp[-(\nu^1)^2-(\nu^2)^2]\,d\nu^1\,d\nu^2. \tag{B.9}$$

(a) The Gauss–Hermite product rule (B.6) with three nodes in each dimension, $Q(3)$, uses one-dimensional nodes and weights given by $\nu_1^h = 0$, $\nu_2^h = \sqrt{\frac{3}{2}}$, $\nu_3^h = -\sqrt{\frac{3}{2}}$, and $\omega_1^h =$

$\frac{2\sqrt{\pi}}{3}$, $\omega_2^h = \omega_3^h = \frac{\sqrt{\pi}}{6}$ for each $h = 1, 2$:

$$\begin{aligned}
E\{G(\epsilon)\} &\approx \frac{1}{\pi}\sum_{j_1=1}^{3}\sum_{j_2=1}^{3}\omega_{j_1}^1\omega_{j_2}^2 G\big(\sqrt{2}\nu_{j_1}^1, \sqrt{2}\nu_{j_2}^2\big) \\
&= \frac{4}{9}G(0,0) + \frac{1}{9}G(0,\sqrt{3}) + \frac{1}{9}G(0,-\sqrt{3}) \\
&\quad + \frac{1}{9}G(\sqrt{3},0) + \frac{1}{36}G(\sqrt{3},-\sqrt{3}) + \frac{1}{36}G(\sqrt{3},-\sqrt{3}) \\
&\quad + \frac{1}{9}G(-\sqrt{3},0) + \frac{1}{36}G(-\sqrt{3},\sqrt{3}) + \frac{1}{36}G(-\sqrt{3},-\sqrt{3}).
\end{aligned}$$

(b) The Gauss–Hermite product rule (B.6) with one node in each dimension, $Q(1)$, uses a node $\nu_1^h = 0$ and a weight $\omega_1^h = \sqrt{\pi}$ for each $h = 1, 2$:

$$E\{G(\epsilon)\} \approx \frac{1}{\pi}\sum_{j_1=1}^{1}\sum_{j_2=1}^{1}\omega_{j_1}^1\omega_{j_2}^2 G\big(\sqrt{2}\nu_{j_1}^1, \sqrt{2}\nu_{j_2}^2\big) = G(0,0).$$

(c) The monomial formula $M1$, given by (B.7), has four nodes:

$$E\{G(\epsilon)\} \approx \frac{1}{4}[G(\sqrt{2},0) + G(-\sqrt{2},0) + G(0,\sqrt{2}) + G(0,-\sqrt{2})].$$

(d) The monomial formula $M2$, given by (B.8), has nine nodes:

$$\begin{aligned}
E\{G(\epsilon)\} &\approx \frac{1}{2}G(0,0) + \frac{1}{16}[G(2,0) + G(-2,0) + G(0,2) + G(0,-2)] \\
&\quad + \frac{1}{16}[G(\sqrt{2},\sqrt{2}) + G(\sqrt{2},-\sqrt{2}) + G(-\sqrt{2},\sqrt{2}) + G(-\sqrt{2},\sqrt{2})].
\end{aligned}$$

Appendix C: Multicountry model

In this section, we provide a formal description of the multicountry model studied in Section 6.6 of the paper. A world economy consists of a finite number of countries N. Each country $h \in \{1, \ldots, N\}$ is populated by a representative consumer. A social planner solves the maximization problem

$$\max_{\{c_t^h, k_{t+1}^h\}_{t=0,\ldots,\infty}^{h=1,\ldots,N}} E_0 \sum_{h=1}^{N}\lambda^h\left[\sum_{t=0}^{\infty}\beta^t u^h(c_t^h)\right] \tag{C.1}$$

subject to the aggregate resource constraint

$$\sum_{h=1}^{N} c_t^h + \sum_{h=1}^{N} k_{t+1}^h = \sum_{h=1}^{N} k_t^h(1-\delta) + \sum_{h=1}^{N} a_t^h A f^h(k_t^h) \tag{C.2}$$

and to the process for the countries' productivity levels

$$\ln a_{t+1}^h = \rho \ln a_t^h + \epsilon_{t+1}^h, \quad h = 1, \dots, N, \tag{C.3}$$

where initial condition $\{k_0^h, a_0^h\}^{h=1,\dots,N}$ is given exogenously, and the productivity shocks follow a multivariate Normal distribution $(\epsilon_{t+1}^1, \dots, \epsilon_{t+1}^N)^\top \sim \mathcal{N}(0_N, \Sigma)$ with $0_N \in \mathbb{R}^N$ being a vector of zero means and $\Sigma \in \mathbb{R}^{N\times N}$ being a variance–covariance matrix. We assume that shocks of different countries are given by $\epsilon_{t+1}^h = \varsigma_{t+1}^h + \varsigma_{t+1}$ and $h = 1, \dots, N$, where $\varsigma_{t+1}^h \sim \mathcal{N}(0, \sigma^2)$ is a country-specific component and $\varsigma_{t+1} \sim \mathcal{N}(0, \sigma^2)$ is a worldwide component. The resulting variance–covariance matrix is

$$\Sigma = \begin{pmatrix} 2\sigma^2 & \cdots & \sigma^2 \\ \cdots & \cdots & \cdots \\ \sigma^2 & \cdots & 2\sigma^2 \end{pmatrix}.$$

In the problem (C.1)–(C.3), E_t denotes conditional expectation; c_t^h, k_t^h, a_t^h, and λ^h are a country h's consumption, capital, productivity level, and welfare weight, respectively; $\beta \in (0, 1)$ is the discount factor; $\delta \in (0, 1]$ is the depreciation rate; A is a normalizing constant in the production function; and $\rho \in (-1, 1)$ is the autocorrelation coefficient. The utility and production functions u^h and f^h, respectively, are strictly increasing, continuously differentiable, and concave. We assume that all countries have identical preferences and technology, that is, $u^h = u$ and $f^h = f$ for all h. Under these assumptions, the planner assigns equal weights $\lambda^h = 1$, and, therefore, equal consumption to all countries, $c_t^h = c_t$ for all $h = 1, \dots, N$.

The solution to the model (C.1)–(C.3) satisfies N Euler equations

$$k_{t+1}^h = E_t\left\{\beta \frac{u'(c_{t+1})}{u'(c_t)}[1 - \delta + a_{t+1}^h A f'(k_{t+1}^h)]k_{t+1}^h\right\}, \quad h = 1, \dots, N, \tag{C.4}$$

where u' and f' are the first derivatives of u and f, respectively.

We approximate the planner's solution in the form of N capital policy functions (45). Note that our approximating functions $\Psi^h(\{k_t^h, a_t^h\}^{h=1,\dots,N}; b^h)$, $h = 1, \dots, N$, are country-specific. Therefore, we treat countries as completely heterogeneous even if they are identical in fundamentals and have identical optimal policy functions. This allows us to assess costs associated with computing solutions to models with heterogeneous preferences and technology.

GSSA, described in Section 2 for the representative–agent model, can be readily adapted to the case of the multicountry model. In the initialization step of Stage 1, we choose an initial guess for the matrix of the coefficients $B \equiv [b^1, \dots, b^N] \in \mathbb{R}^{(n+1)\times N}$ in the assumed approximating functions $\Psi^h(\{k_t^h, a_t^h\}^{h=1,\dots,N}; b^h)$, $h = 1, \dots, N$. In Step 1, at iteration p, we use a matrix $B^{(p)}$ to simulate the model T periods forward to obtain $\{k_{t+1}^h\}_{t=0,\dots,T}^{h=1,\dots,N}$ and calculate the average consumption $\{c_t\}_{t=0,\dots,T}$ using the resource constraint (C.2). In Step 2, we calculate the conditional expectation in (C.4) using a selected integration method to obtain $\{y_t^h\}_{t=0,\dots,T-1}^{h=1,\dots,N}$. In Step 4, we run N regressions $y_t^h = \Psi^h(\{k_t^h, a_t^h\}^{h=1,\dots,N}; b^h) + \varepsilon_t^h$ to obtain a new matrix of the coefficients $\widehat{B} = [\widehat{b}^1, \dots, \widehat{b}^N]$; as in the representative–agent case, we assume that Ψ^h is linear in b^h,

which leads to a linear regression model $y^h = Xb^h + \varepsilon^h$, where $y^h \equiv (y_0^h, \ldots, y_{T-1}^h)^\top \in \mathbb{R}^T$, $\varepsilon^h \equiv (\varepsilon_0^h, \ldots, \varepsilon_{T-1}^h)^\top \in \mathbb{R}^T$, and $X \in \mathbb{R}^{T\times(n+1)}$ is a matrix of explanatory variables constructed with the basis functions of the state variables. Finally, in Step 4, we update the coefficients B using fixed-point iteration, $B^{(p+1)} = (1-\xi)B^{(p)} + \xi\widehat{B}$. In Stage 2, we evaluate the Euler equation errors on a simulation of $T^{\text{test}} = 10{,}200$ observations using a high-quality integration method: for $N \leq 20$, we use the monomial rule $M2$ and for $N > 20$, we use the monomial rule $M1$. To solve the model, we assume $u(c_t) = \ln c_t$, $f(k_t) = k_t^\alpha$ with $\alpha = 0.36$, $\beta = 0.99$, $\delta = 0.025$, $\rho = 0.95$, and $\sigma = 0.01$.

References

Judd, K. (1998), *Numerical Methods in Economics*. MIT Press, Cambridge, Massachusetts.

Levenberg, K. (1944), "A method for the solution of certain nonlinear problems in least squares." *Quarterly Applied Mathematics*, 4, 164–168.

Marcet, A. (1988), "Solving non-linear models by parameterizing expectations." Unpublished manuscript, Carnegie Mellon University, Graduate School of Industrial Administration.

Marquardt, D. (1963), "An algorithm for least-squares estimation of nonlinear parameters." *Journal of Society for Industrial Applied Mathematics*, 11, 431–441.

Stroud, A. (1971), *Approximate Integration of Multiple Integrals*. Prentice Hall, Englewood Cliffs, New Jersey.