

The method of endogenous gridpoints for solving dynamic stochastic optimization problems

Christopher D. Carroll

Department of Economics, The Johns Hopkins University, Baltimore MD, 21218-2685, USA

Received 21 December 2004; received in revised form 29 July 2005; accepted 6 September 2005

Abstract

This paper introduces a solution method for numerical dynamic stochastic optimization problems that avoids rootfinding operations. The idea is applicable to many microeconomic and macroeconomic problems, including life cycle, buffer-stock, and stochastic growth problems. Software is provided.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Dynamic optimization; Precautionary saving; Stochastic growth model; Endogenous gridpoints; Liquidity constraints

JEL classification: C6; D9; E2

1. The Problem

Consider a consumer whose goal is to maximize discounted utility from consumption

$$\max \sum_{s=t}^T \beta^{s-t} u(C_s) \quad (1)$$

for a CRRA utility function $u(C) = C^{1-\rho}/(1-\rho)$.¹

E-mail address: ccarroll@jhu.edu.

URL: <http://econ.jhu.edu/people/ccarroll>. Software archive at <http://econ.jhu.edu/people/ccarroll/EndogenousArchive.zip>.

¹ Putting leisure in the utility function is straightforward but would distract from the paper's point.

The consumer's problem will be specialized below to two cases: a standard micro-economic problem with uninsurable idiosyncratic shocks to labor income, and a standard representative agent problem with shocks to aggregate productivity (the 'micro' and the 'macro' models).²

The consumer's initial condition is defined by two state variables: M_t is 'market resources' (macro interpretation: capital plus current output) or 'cash-on-hand' (micro interpretation: net worth plus current income), while P_t is permanent labor productivity in both interpretations.

The transition process for M_t is broken up, for convenience of analysis, into three steps. Assets at the end of the period are market resources minus consumption, equal to

$$A_t = M_t - C_t, \quad (2)$$

and capital at the beginning of the next period is what remains after a depreciation factor τ is applied,

$$K_{t+1} = A_t \tau, \quad (3)$$

where $\tau = (1 - \delta)$ in the usual macro notation and $\tau = 1$ in the micro interpretation.

The final step can be thought of as the transition from the beginning of period $t + 1$, when capital K_{t+1} but has not yet been used to produce output, and the middle of that period, when output has been produced and incorporated into resources:

$$M_{t+1} = \overbrace{e_{t+1} \Theta_{t+1} P_{t+1}}^{\equiv L_{t+1}} \mathcal{W}_{t+1} + K_{t+1} \mathcal{R}_{t+1} \quad (4)$$

where \mathcal{W}_{t+1} is the wage rate; Θ_{t+1} is an iid transitory shock (e.g., unemployment) normalized to satisfy $E_t[\Theta_{t+n}] = 1 \forall n > 0$ (usually $\Theta_t = 1 \forall t$ in the macro interpretation); and e_t indicates labor effort (or labor supply), which for purposes of this paper is fixed at $e_t = 1$, but in general could be allowed to vary. The **disarticulation** of the flow of income into labor and capital components is useful in thinking separately about the effects of productivity growth (captured by ΘP) and capital accumulation (K).

Permanent labor productivity (in either interpretation) evolves according to

$$P_{t+1} = G_{t+1} P_t \Psi_{t+1} \quad (5)$$

for a permanent shock that satisfies $E_t[\Psi_{t+n}] = 1 \forall n > 0$ and G_t is exogenous and perfectly predictable (see below for varying interpretations of G).

Defining lower case variables as the upper-case variable scaled by the level of permanent labor productivity, e.g. $a_t = A_t / P_t$, we have

$$a_t = m_t - c_t \quad (6)$$

while with a bit of algebra the state transition becomes

$$m_{t+1} = \underbrace{e_t \Theta_{t+1}}_{\equiv l_{t+1}} \mathcal{W}_{t+1} + \underbrace{(a_t \tau / G_{t+1} \Psi_{t+1})}_{\equiv k_{t+1}} \mathcal{R}_{t+1}. \quad (7)$$

The interest and wage factors are assumed not to depend on anything other than capital and productive labor input; together with the iid assumption about the structure of the shocks, this implies that the problem has a Bellman equation representation (henceforth **boldface** indicates functions)

$$\mathbf{V}_t(M_t, P_t) = \max_{C_t} \{u(C_t) + \beta E_t[\mathbf{V}_{t+1}(M_{t+1}, P_{t+1})]\} \quad (8)$$

subject to the transition equations.

² Different aspects of the setup of the problem will strike micro and macroeconomists as **peculiar**; with patience, it should become clear how the problem as specified can be transformed into more familiar forms.

Defining $\Lambda_{t+1} \equiv G_{t+1} \Psi_{t+1}$, consider the related problem

$$\mathbf{v}_t(m_t) = \max_{c_t} \left\{ u(c_t) + \beta E_t \left[\Lambda_{t+1}^{1-\rho} \mathbf{v}_{t+1} \left(\overbrace{\mathcal{W}_{t+1} l_{t+1} + \mathcal{R}_{t+1} \underbrace{a_t \mathfrak{T} / \Lambda_{t+1}}_{k_{t+1}}}^{=m_{t+1}}} \right) \right] \right\}. \quad (9)$$

Assume that there is some last period T in which

$$\mathbf{V}_T(M_T, P_T) = P_T^{1-\rho} \mathbf{v}_T(M_T/P_T) \quad (10)$$

for some well-behaved \mathbf{v}_T (we will be more specific about the terminal value function below). In this case it is easy to show that the solution to the ‘normalized’ problem defined by Eq. (9) yields the solution to the original problem via $\mathbf{V}_t = P_t^{1-\rho} \mathbf{v}_t$ for any $t < T$.³

Now define an end-of-period value function ‘Gothic v’ as

$$\mathbf{v}_t(a_t) = \beta E_t \left[\Lambda_{t+1}^{1-\rho} \mathbf{v}_{t+1} (\mathcal{W}_{t+1} l_{t+1} + \mathfrak{T} \mathcal{R}_{t+1} a_t / \Lambda_{t+1}) \right] \quad (11)$$

with derivative

$$\begin{aligned} \mathbf{v}_t^a(a_t) &= \beta E_t \left[\Lambda_{t+1}^{1-\rho} \mathbf{v}_{t+1}^m (\mathcal{W}_{t+1} l_{t+1} + \mathcal{R}_{t+1} a_t \mathfrak{T} / \Lambda_{t+1}) \mathcal{R}_{t+1} \mathfrak{T} / \Lambda_{t+1} \right] \\ &= \mathfrak{T} \beta E_t \left[\Lambda_{t+1}^{-\rho} \mathbf{v}_{t+1}^m (\mathcal{W}_{t+1} l_{t+1} + \mathcal{R}_{t+1} a_t \mathfrak{T} / \Lambda_{t+1}) \mathcal{R}_{t+1} \right] \end{aligned} \quad (12)$$

and Eqs. (11) and (6) imply that Eq. (9) can be rewritten using \mathbf{v}_t as

$$\mathbf{v}_t(m_t) = \max_{\{a_t\}} \{u(m_t - a_t) + \mathcal{V}_t(a_t)\}, \quad (13)$$

and the envelope theorem can be applied

$$\mathbf{v}_t^m(m_t) = u'(c_t) \quad (14)$$

while the first order condition yields the Euler equation

$$u'(c_t) = \mathbf{V}_t^a(a_t) = \mathfrak{T} \beta E_t [u'(\Lambda_{t+1} c_{t+1}) \mathcal{R}_{t+1}]. \quad (15)$$

2. Recursion

Generically, problems like this can be solved by specifying a final-period decision rule \mathbf{c}_T and a procedure for recursion (obtaining \mathbf{c}_t from \mathbf{c}_{t+1}). Here we specify the recursion; below we specify choices for the terminal decision rule.

2.1. A standard solution method

The absence of a closed-form solution means that optimal decision functions (e.g. the consumption function) must be constructed by calculating their values at a finite grid of possible values of the state variables. Call some ordered set of such values $\mu_i \in \vec{\mu} \equiv \{\mu_1, \mu_2, \dots, \mu_I\}$.

³ See Carroll (2004) for a proof.

With c_{t+1} in hand, the usual solution procedure is to specify a $\vec{\mu}$ and, for each element μ_i , to use a numerical rootfinding routine to find the χ_i that satisfies Eq. (15),

$$u'(\chi_i) = \mathcal{V}_a^t(\mu_i - \chi_i). \quad (16)$$

The points $\{\mu_i, \chi_i\}$ are then used to construct an interpolating approximation to c_t . (Choice of interpolation method is separable from the point of this paper; see Judd (1998) for a discussion of choices). Given the interpolated c_t function the solution for earlier periods is found by recursion.

One of the most computationally burdensome steps in this approach is the numerical solution of Eq. (16) for each specified state gridpoint. Even if efficient methods are used for constructing the expectations (cf. the parameterized-expectations method of denHaan and Marcet (1990)) and shrewd choices are made for the points to include in $\vec{\mu}$, for each gridpoint a numerical rootfinding operation still must evaluate a substantial number of candidate values for the control variable before finding values that satisfy Eq. (16) to an acceptable degree of precision.

2.2. Endogenous gridpoints solution method

This paper's key contribution is to introduce an alternative approach that does not require numerical rootfinding. The trick is to begin with *end-of-period* assets a_t and to use the end-of-period marginal value function v_t^a , the first order condition, and the budget constraint to construct the unique values of middle-of-period m_t generated by those at values.

Specifically, define an exogenous, time-invariant ordered set of values of a_t collected in $\alpha \in \vec{\alpha} \equiv \{\alpha_1, \alpha_2, \dots, \alpha_I\}$. For each end-of-period state α_i the marginal value $V_t^a(\alpha_i)$ is easy to calculate; inverting the consumption first order condition, the α 's generate

$$\chi_i = u'^{-1}(V_t^a(\alpha_i)). \quad (17)$$

Note that the budget constraint implies that

$$u_i = \alpha_i + \chi_i. \quad (18)$$

We now have a collection of $\{\mu_i, \chi_i\}$ pairs in hand and can interpolate as before to generate an approximation to c_t . This completes the recursion.

The key distinction between this approach and the standard one is that the gridpoints for the policy functions are not predetermined; instead they are endogenously generated from a predetermined grid of values of end-of-period assets (hence the method's name). One reason the method is efficient is that expectations are never computed for any grid-point not used in the final interpolating function; the standard method may compute expectations for many unused gridpoints.

3. Macro specialization

We first specialize to a macroeconomic stochastic growth model. Assuming aggregate production is Cobb–Douglas in capital and labor $F(K, P) = K^\epsilon P^{1-\epsilon}$, after normalizing by productivity P (and

assuming a constant value G for the labor productivity growth factor), under the usual assumptions of perfect competition, etc., if there is no aggregate transitory shock ($\Theta_{t+1}=1$) we have

$$\mathcal{R}_{t+1} = 1 + \varepsilon k_{t+1}^{\varepsilon-1} \quad (19)$$

$$\mathcal{W}_{t+1} = (1 - \varepsilon)k_{t+1}^{\varepsilon} \quad (20)$$

and market resources are the sum of capital and production,

$$m_{t+1} = k_{t+1}\mathcal{R}_{t+1} + \mathcal{W}_{t+1} \quad (21)$$

$$= k_{t+1} + k_{t+1}^{\varepsilon}. \quad (22)$$

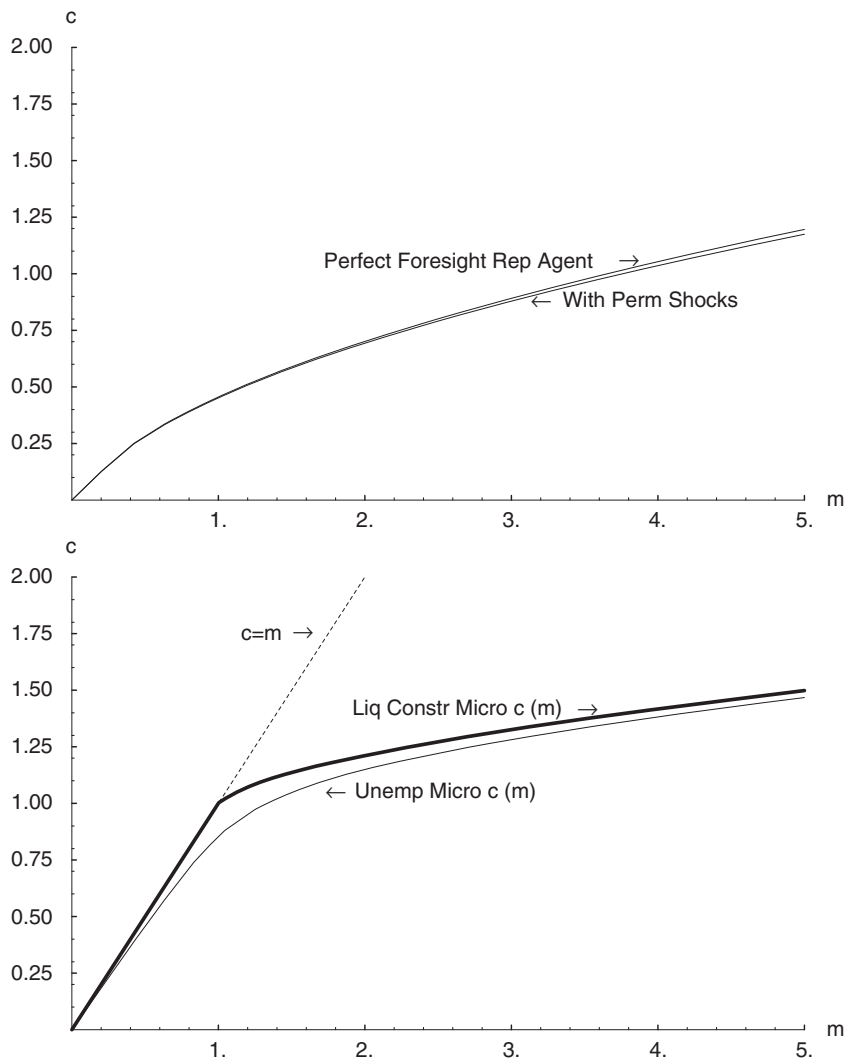


Fig. 1. Macro- and micro-consumption functions.

Table 1
Parameter values

<i>Parameters common to all models</i>		
ρ	2	Relative risk aversion
β	0.96	Annual discount factor
e	1	Labor supply/effort (fixed)
$\vec{\Psi}$	{0.90, 1.00, 1.10}	Permanent shock realizations
$Pr(\vec{\Psi})$	{0.25, 0.50, 0.25}	Permanent shock probabilities
<i>Macro model parameters</i>		
τ	0.90	Depreciation factor
G	1.01	Exogenous aggregate productivity growth factor
ε	0.36	Capital share in production
<i>Micro model parameters</i>		
τ	1	Depreciation factor
G	1.03	Trend individual wage growth factor
R	1.04	Real interest rate
W	1.00	Wage rate
$\vec{\Xi}$	{0.90, 1.00, 1.10}	Transitory shock realizations for employed
$Pr(\vec{\Xi} \Theta>0)$	{0.25, 0.50, 0.25}	Transitory shock probabilities for employed
<i>Parameter unique to unemployment model</i>		
ϕ	0.005	Probability of unemployment spell

We specify the terminal consumption function as

$$c_T(m) = m, \quad (23)$$

which is very far from the converged infinite horizon consumption rule, but easy to verify as satisfying the assumption (10) imposed earlier. More efficient choices are available, but for our purposes simplicity trumps efficiency.

An arbitrary specification of the process for permanent productivity shocks is a three point distribution defined by $\vec{\Psi} = \{0.9, 1.0, 1.1\}$ with probabilities $Pr(\vec{\Psi}) = \{0.25, 0.50, 0.25\}$.⁴

The top panel of Fig. 1 plots the converged consumption function that emerges from this solution method for the benchmark set of parameter values specified in Table 1, along with the consumption function for the standard perfect foresight version of the model ($\vec{\Psi} = Pr(\vec{\Psi}) = \{1\}$).

4. Micro specialization

In the microeconomic literature, the usual approach is to take aggregate interest and wage rates as exogenous, and to focus on transitory (Θ) and permanent (Ψ) shocks to idiosyncratic labor productivity.

⁴ With careful choice of points and weights, small-dimensional discrete representations like this do a good job of approximating commonly used continuous distributions like a log normal, cf. Judd (1998). An empirically realistic choice would have a much lower variance than the specification here.

We again start the recursion with $c_T(m)=m$, and the permanent shocks are retained exactly as specified for the macro problem.⁵

4.1. Life cycle models

Life cycle models specify a stereotypical pattern of lifetime income growth defined by G_t where t is age rather than time and T is the maximum possible lifespan;⁶ mortality uncertainty can be accommodated by age-varying values of β .

4.2. Buffer stock models

If R , W , G and β are constant, $\tau=1$, and the impatience condition

$$R\beta E[(G\Psi)^{-\rho}] < 1 \quad (24)$$

holds, Deaton (1991) and Carroll (2004) show that the problem defines a contraction mapping so that the consumption functions defined by the problem converge from any well-behaved initial starting function $c_T(m)$; the converged function is defined as

$$c(m) = \lim_{n \rightarrow \infty} c_{T-n}(m). \quad (25)$$

We solve for the converged consumption function for two versions.

4.2.1. Version with unemployment

Assume that in future periods there is a small probability \mathcal{P} that income will be zero (corresponding to a substantial spell of unemployment):

$$\Theta_{t+1} = \begin{cases} 0 & \text{with probability } p > 0 \\ \bar{\varepsilon}_{t+1}/(1-p) & \text{with probability } (1-p) \end{cases} \quad (26)$$

where $\bar{\varepsilon} = \{0.9, 1.0, 1.1\}$ and $Pr(\bar{\varepsilon}\Theta > 0) = \{0.25, 0.50, 0.25\}$ (the same structure of non-unemployment transitory shocks as for the permanent shocks).

Carroll (2004) shows that in this model,

$$\lim_{m_t \rightarrow 0} c_t(m_t) = 0. \quad (27)$$

This implies that the minimum value in $\vec{\alpha}$ should be $\alpha_1=0$, which will generate $\{\mu_1, \chi_1\} = \{0., 0.\}$ as the first point in the set of interpolating points. The resulting converged $c(m)$ is shown as the thin solid

⁵ An empirically realistic calibration for micro data would exhibit a permanent variance perhaps 100 times greater than an appropriate macro calibration; but appropriate calibration is not the point of this paper.

⁶ This is the context in which the assumption that $c_T(m)=m$ actually makes economic sense, as distinct from merely providing a convenient starting point for recursion.

locus in the bottom panel of Fig. 1; see the software for details of how the remaining values in $\vec{\alpha}$ were chosen.

4.2.2. Version with liquidity constraints

Microeconomic models often include a liquidity constraint in addition to the usual transition equations, and capturing the constraint often induces much additional code.

Dealing with a liquidity constraint using the method of endogenous gridpoints is simple. The key observation is that when the constraint is on the cusp of binding, the marginal value of consumption is equal to the marginal value of saving exactly zero (assuming the constraint is of the form that requires a to be nonnegative; generalization to more elaborate kinds of constraints is straightforward). If the first value in the ordered set $\vec{\alpha}$ is $\alpha_1=0$, then the method will produce

$$\chi_1 = \mu_1 = u'^{-1}(\mathbf{V}_t^a(0)), \quad (28)$$

and if we define $\hat{c}_t(m)$ as the function produced by interpolation among the points generated by $\vec{\alpha}$, the consumption function imposing the constraint will be

$$c_t(m) = \min(m, \hat{c}_t(m)). \quad (29)$$

If the consumption function is defined as a piecewise linear spline interpolation among the $\{\mu, \chi\}$ points, the constraint can be handled simply by adding the point $\{\mu_0, \chi_0\} = \{0, 0\}$ to the set of points that constitute the interpolation data.

The converged solution is shown as the bold locus in the bottom panel of Fig. 1.

5. Conclusion

The method of endogenous gridpoints can be extended to problems with multiple state variables and multiple controls, e.g. a micro consumer with a portfolio choice problem, or a labor supply decision; or a macro consumer with a utility function that exhibits habit formation (see Carroll (2000) for examples). The method is useful both because it is simpler than the standard method and because it reduces computational demands.

Acknowledgements

I am grateful to Nicola Fuchs-Schündeln, Wouter den Haan, Michael Haliassos, Ken Judd, Albert Marcet, Dimitri Mavridis, Victor Rios-Rull, John Rust, Eric Young, and participants in the 2002 meetings of the Society for Computational Economics for discussions of the literature and existing solution techniques. Excellent research assistance has been provided by Zhou Lu and Marc Chan. I am especially grateful to Michael Reiter for improvements to the Matlab code. All errors are my own.

References

- Carroll, Christopher D., 2000. Lecture notes on solving microeconomic dynamic stochastic optimization problems Manuscript, Johns Hopkins University. Available at <http://econ.jhu.edu/people/ccarroll/solvingmicrodsops.pdf>.

- Carroll, Christopher D., 2004. Theoretical foundations of buffer stock saving NBER Working Paper Number 10867. <http://econ.jhu.edu/people/ccarroll/BufferStockProofsNew.pdf>.
- Deaton, Angus S., 1991. Saving and liquidity constraints. *Econometrica* 59, 1221–1248.
- denHaan, Wouter J., Marcet, Albert, 1990. Solving the stochastic growth model by parameterizing expectations. *Journal of Business and Economic Statistics* 8 (1), 31–34. available at. <http://ideas.repec.org/a/bes/jnlbes/v8y1990i1p31-34.html>.
- Judd, Kenneth L., 1998. *Numerical Methods in Economics*. The MIT Press, Cambridge, MA.