Discretizing Continuous-Time Operators with Finite Differences

Jesse Perla
UBC

September 18, 2017

Here, we expand on details for how to discretize linear and nonlinear operators for use in other methods.[1] Most of the notes will focus on time-homogeneity examples.[2]

General Notation   The following expressions are used in the document:

- Given an operator (or infinitesimal generator) associated with a particular stochastic process, $\mathcal{A}$.[3] The purpose of these notes is to discretize $\mathcal{A}$ on this grid using using finite differences. Crucially, it is necessary to put boundary-values into the discretized operator as well.

- Where appropriate, we will define the adjoint of the operator $\mathcal{A}$ as $\mathcal{A}^*$. This is useful when trying to solve for the evolution of distributions (rather than solving for the value functions).

- For a given variable q, define the notation $\mu^- \equiv \min\{q, 0\}$ and $q^+ \equiv \max\{q, 0\}$, which will be useful for defining finite-differences with an upwind scheme.[4] This can apply to vectors as well. For example, $q_i^- = q_i$ if $q_i < 0$ and 0 if $q_i > 0$, and $q_i^- \equiv \left\{ q_i^- \right\}_{i=1}^I$.

- Finally, derivatives are denoted by the operator $\boldsymbol{\partial}$ and univariate derivatives such as $\boldsymbol{\partial}_x v(x) \equiv v'(x)$.

Grids   Start with the simplification of a univariate function of x,

- In the univariate case with I points, $\{x_i\}_{i=1}^I$ with $x_1 = \underline{x}$ and $x_I = \bar{x}$ when $x \in [\underline{x}, \bar{x}]$. After discretizing, we will denote the grid with the variable name, i.e. $x \equiv \{x_i\}_{i=1}^I$

- For the grid x, denote the forward and backward distance between grid points at node i as,

$$\Delta_{i,+} \equiv x_{i+1} - x_i \tag{1}$$
$$\Delta_{i,-} \equiv x_i - x_{i-1} \tag{2}$$

In the simple case of a uniform grid, $\Delta \equiv \Delta_{i,+} = \Delta_{i,-} = x_{i+1}$ for all i

- When we discreteize a function, use the function name without arguments to denote the vector. i.e. $v(x)$ discretized on a grid $\{x_i\}_{i=1}^I$ is $v \equiv \{v(x_i)\}_{i=1}^I \in \mathbb{R}^I$

---

[1]These notes expand on Ben Moll's superb notes in http://www.princeton.edu/~moll/HACTproject/.

[2]For more background on differential operators, see https://en.wikipedia.org/wiki/Differential_operator.

[3]See https://en.wikipedia.org/wiki/Infinitesimal_generator_(stochastic_processes) for some formulas and interpretation for diffusions, and https://en.wikipedia.org/wiki/Transition_rate_matrix for the generator of continuous-time Markov chains.

[4]For more details on the notation and the upwind scheme, see http://www.princeton.edu/~moll/HACTproject/HACT_Numerical_Appendix.pdf.

1 Time-Independent Univariate Diffusions

Assume that the time variable does not directly enter into anything important in the problem (e.g. drifts, payoffs, stopping values, boundary values, or discount rates). This is also assuming that the drift, variance, and boundary values are not a (direct) function of any v_i.[5] Otherwise, the operator would be nonlinear. In the case of a linear $\mathcal{A}$, the resulting discretized operator will be a (heavily sparse) matrix $A \in \mathbb{R}^{I \times I}$, which is especially convenient to solve.

1.1 Stochastic Process and Boundary Values

Take a diffusion process for x_t according to the following stochastic difference equation,

$$\mathrm{d}x_t = \mu(x_t)\mathrm{d}t + \sigma(x_t)\mathrm{d}\mathbb{W}_t \tag{3}$$

where $\mathbb{W}_t$ is Brownian motion and $x \in (\underline{x}, \bar{x})$ where $-\infty \leq \underline{x} < \bar{x} \leq \infty$.[6] The functions $\mu(\cdot)$ and $\sigma(\cdot)$ are left general, but it is essential that they are time-homogeneous.

The partial differential operator (infinitesimal generator) associated with the stochastic process in (3) is[7]

$$\mathcal{A} \equiv \mu(x)\boldsymbol{\partial}_x + \frac{\sigma(x)^2}{2}\boldsymbol{\partial}_{xx} \tag{4}$$

We will consider several types of boundary values for an operator on a generic function $v(x)$. Because the process is univariate, we require boundary conditions at both sides of x, and assume that the boundary and $\mu(\cdot), \sigma(x)^2$ are independent of time. Given that the operator is on a function $v(x)$, consider:

- **Absorbing Barrier:** $v(\underline{x}) = \underline{v}$ and/or $v(\bar{x}) = \bar{v}$ for constants $\underline{v}$ and/or $\bar{v}$. This is called a "Dirichlet" boundary condition in PDEs.

- **Reflecting Barrier:** $\boldsymbol{\partial}_x v(\underline{x}) = v'(\underline{x}) = 0$ and/or $v'(\bar{x}) = 0$. This is called a "Neumann" boundary condition in PDEs.

- **Constrained Process:** In many cases, bounds on the variable are not essential to the process, but really just necessary to discretize. In those cases, adding in an artificial "reflecting barrier" at some large $\bar{x}$ (or small $\bar{x}$) seems to be the best approach to constraining the process for numerical methods.[8] The hope is that the introduction of this boundary condition would have little impact on the solution except very close to the artificial boundary.

- **Transversality:** Problem specific? But the hope is that using a "Constrained Process" with a reflecting barrier converges to the correct solution in most of the domain.

To distinguish between the boundary values we will consider, define

$$R_{\underline{x}} \equiv \begin{cases} 1 & \text{if } x \text{ is reflected at } \underline{x} \\ 0 & \text{if there is a absorbing barrier with value } \underline{v} \end{cases} \tag{5}$$

$$R_{\bar{x}} \equiv \begin{cases} 1 & \text{if } x \text{ is reflected at } \bar{x} \\ 0 & \text{if there is a absorbing barrier with value } \bar{v} \end{cases} \tag{6}$$

---

[5]However, in the case of controlled-diffusions a numerical method is to use the last iterations v_i^{n-1} as part of the μ_i^n definition for the equations in v_i^n. This process has a linear $\mathcal{A}$ at any given iteration.

[6]We are being a little sloppy with notation x_t being exactly at the bounds because it is a measure 0 event in these setups.

[7]When time enters payoffs, stopping values, etc. the generator is instead on a function of both t an x. See Section 2

[8]See Kushner and Dupuis (2001) section 1.4, 5.7, and 11.1 for a discussion.

What Boundary Value to Assume? In general, a reflection for a large $\bar{x}$ and small $\underline{x}$ is the most appropriate. The exceptions would tend to have direct economics associated with a death or uncontrolled stopping.[9]

1.2   Upwind Finite Difference Discretization for A with a Uniform Grid

Denote the vector of drifts and variances as $\mu \equiv \{\mu(x_i)\}_{i=1}^I$ and $\sigma^2 \equiv \{\sigma(x_i)^2\}_{i=1}^I$. For a uniform grid with $\Delta \equiv x_{i+1} - x_i$. Define the vectors, $X, Y, Z \in \mathbb{R}^I$ such that,

$$X = -\frac{\mu^-}{\Delta} + \frac{\sigma^2}{2 \times \Delta^2} \tag{7}$$

$$Y = -\frac{\mu^+}{\Delta} + \frac{\mu^-}{\Delta} - \frac{\sigma^2}{\Delta^2} \tag{8}$$

$$Z = \frac{\mu^+}{\Delta} + \frac{\sigma^2}{2 \times \Delta^2} \tag{9}$$

With these, the matrix A is constructed as.

$$A \equiv \begin{bmatrix} Y_1 + R_{\underline{x}}X_1 & Z_1 & 0 & \cdots & \cdots & \cdots & 0 \\ X_2 & Y_2 & Z_2 & 0 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & X_{I-1} & Y_{I-1} & Z_{I-1} \\ 0 & \cdots & \cdots & \cdots & 0 & X_I & Y_I + R_{\bar{x}}Z_I \end{bmatrix} \in \mathbb{R}^{I \times I} \tag{10}$$

$$b \equiv \begin{bmatrix} (1 - R_{\underline{x}})X_1\underline{v} \\ 0 \\ \vdots \\ 0 \\ (1 - R_{\bar{x}})Z_I\bar{v} \end{bmatrix}^T \in \mathbb{R}^I \tag{11}$$

where,

$$\mathcal{A}v(\{x_i\}) \approx Av + b \tag{12}$$

Note that the special cases for the 1st and Ith row correspond to the boundary values and are adjusted depending on the type of boundary value. Also, note that the b vector is all zeros if the boundary values are reflective and/or the matched value $\bar{v}$ or $\underline{v}$ is 0).

Intensity Matrix of a Continuous-Time Markov Chain Note that with reflecting barriers on both sides, the sum for each row of A is $X_i + Y_i + Z_i = 0$ for all $i = 1, \ldots I$ from (7) to (9), and it can be shown that the diagonal is always negative. Hence, we interpret the A matrix fulfills the requirements of an intensity matrix (or infinitesimal generator matrix) of a continuous-time Markov chain.[10] Furthermore, the Markov-chain has the convenient property that jumps are only between adjacent states (i.e. if at x_i then can only jump to x_{i-1} and x_{i+1}), which is extremely sparse.

---

[9]For example, even if you have firms optimally exiting, you could use a small reflection at $\bar{x}$ for the variational methods in `optimal_stopping.pdf` as long as $\bar{x}$ is below any binding level.

[10]See `https://en.wikipedia.org/wiki/Transition_rate_matrix` and `https://en.wikipedia.org/wiki/Markov_chain`.

When the boundary value is absorbing at $\underline{x}$ or $\bar{x}$, the interpretation is a little trickier since the sum of all elements is not 0. Essentially, we need to take into account "death" since the agents are no longer reflected within the boundaries at all times. This is the reason for a non-zero b matrix.

Interior of A: To better understand the construction of A and b, look at individual rows of A with the ODE.

In the interior $(1 < i < I)$, the discretization of (4)

$$\mathcal{A}v(x_i) \approx \underbrace{\frac{v_i - v_{i-1}}{\Delta}\mu_i^- + \frac{v_{i+1} - v_i}{\Delta}\mu_i^+}_{\text{Upwind Scheme}} + \frac{\sigma_i^2}{2}\frac{v_{i+1} - 2v_i + v_{i-1}}{\Delta^2} \tag{13}$$

The upwind scheme chooses either forward or backward differences, depending on the sign of the drift. Collecting terms, we see the derivation for the definitions in (7) to (9)

$$= \underbrace{\left(-\frac{\mu_i^-}{\Delta} + \frac{\sigma_i^2}{2\Delta^2}\right)}_{\equiv X_i} v_{i-1} + \underbrace{\left(\frac{\mu_i^-}{\Delta} - \frac{\mu_i^+}{\Delta} - \frac{\sigma_i^2}{\Delta^2}\right)}_{\equiv Y_i} v_i + \underbrace{\left(\frac{\mu_i^+}{\Delta} + \frac{\sigma_i^2}{2\Delta^2}\right)}_{\equiv Z_i} v_{i+1} \tag{14}$$

Boundary Value at $\underline{x}$: As $i = 1$, given the "ghost node" x_0, the discretized operator from (14) is

$$\mathcal{A}v(x_1) \approx \left(-\frac{\mu_1^-}{\Delta} + \frac{\sigma_1^2}{2\Delta^2}\right) v_0 + \left(\frac{\mu_1^-}{\Delta} - \frac{\mu_1^+}{\Delta} - \frac{\sigma_1^2}{\Delta^2}\right) v_1 + \left(\frac{\mu_1^+}{\Delta} + \frac{\sigma_1^2}{2\Delta^2}\right) v_2 \tag{15}$$

In the case of the boundary value $v(\underline{x}) = \underline{v}$, subsitute for $v_0 = \underline{v}$ to find,

$$\approx \overbrace{\underbrace{\left(-\frac{\mu_1^-}{\Delta} + \frac{\sigma_1^2}{2\Delta^2}\right)}_{\equiv b_1}}^{\equiv X_1} \underline{v} + \underbrace{\left(\frac{\mu_1^-}{\Delta} - \frac{\mu_1^+}{\Delta} - \frac{\sigma_1^2}{\Delta^2}\right)}_{\equiv Y_1} v_1 + \underbrace{\left(\frac{\mu_1^+}{\Delta} + \frac{\sigma_1^2}{2\Delta^2}\right)}_{\equiv Z_1} v_2 \tag{16}$$

Alternatively, if the boundary value is $v'(\underline{x}) = 0$, take (15) and use $v'(\underline{x}) \approx \frac{v_1 - v_0}{\Delta} = 0$, $\implies v_1 = v_0$, so

$$\mathcal{A}v(x_1) \approx \left(\underbrace{\left(-\frac{\mu_1^-}{\Delta} + \frac{\sigma_1^2}{2\Delta^2}\right)}_{\equiv X_1} + \underbrace{\left(\frac{\mu_1^-}{\Delta} - \frac{\mu_1^+}{\Delta} - \frac{\sigma_1^2}{\Delta^2}\right)}_{\equiv Y_1}\right) v_1 + \underbrace{\left(\frac{\mu_1^+}{\Delta} + \frac{\sigma_1^2}{2\Delta^2}\right)}_{\equiv Z_1} v_2 \tag{17}$$

Boundary Value at $\bar{x}$: As $i = I$, given the "ghost node" x_{I+1}, from (14)

$$\mathcal{A}v(x_I) \approx \left(-\frac{\mu_I^-}{\Delta} + \frac{\sigma_I^2}{2\Delta^2}\right) v_{I-1} + \left(\frac{\mu_I^-}{\Delta} - \frac{\mu_I^+}{\Delta} - \frac{\sigma_I^2}{\Delta^2}\right) v_I + \left(\frac{\mu_I^+}{\Delta} + \frac{\sigma_I^2}{2\Delta^2}\right) v_{I+1} \tag{18}$$

For the absorbing barrier, substitute for $v(x_{I+1}) = \bar{v}$,

$$\mathcal{A}v(x_I) \approx \underbrace{\left(-\frac{\mu_I^-}{\Delta} + \frac{\sigma_I^2}{2\Delta^2}\right)}_{\equiv X_I} v_{I-1} + \underbrace{\left(\frac{\mu_I^-}{\Delta} - \frac{\mu_I^+}{\Delta} - \frac{\sigma_I^2}{\Delta^2}\right)}_{\equiv Y_I} v_I + \overbrace{\underbrace{\left(\frac{\mu_I^+}{\Delta} + \frac{\sigma_I^2}{2\Delta^2}\right)}_{\equiv b_I}}^{\equiv Z_I} \bar{v} \tag{19}$$

4

For a reflecting barrier, the boundary value $v'(\bar{x}) \approx \frac{v_{I+1}-v_I}{\Delta} = 0, \implies v_{I+1} = v_I,$

$$\mathcal{A}v(x_I) \approx \underbrace{\left(-\frac{\mu_I^-}{\Delta} + \frac{\sigma_I^2}{2\Delta^2}\right)}_{\equiv X_I} v_{I-1} + \left(\underbrace{\left(\frac{\mu_I^-}{\Delta} - \frac{\mu_I^+}{\Delta} - \frac{\sigma_I^2}{\Delta^2}\right)}_{\equiv Y_I} + \underbrace{\left(\frac{\mu_I^+}{\Delta} + \frac{\sigma_I^2}{2\Delta^2}\right)}_{\equiv Z_I}\right) v_I \tag{20}$$

Simple Case with $\mu(x) < 0$ and Reflecting Barrier In the special case of $\mu(\bar{x}) < 0$, the upwind direction is trivial: $\mu^- = \mu$ and $\mu^+ = 0$. Specializing (7) to (9),

$$X = -\frac{\mu}{\Delta} + \frac{\sigma^2}{2 \times \Delta^2} \tag{21}$$

$$Y = \frac{\mu}{\Delta} - \frac{\sigma^2}{\Delta^2} \tag{22}$$

$$Z = \frac{\sigma^2}{2 \times \Delta^2} \tag{23}$$

and the formula for the discretized operator is,

$$\mathcal{A}v(x_i) = \frac{v_i - v_{i-1}}{\Delta}\mu_i + \frac{\sigma_i^2}{2}\frac{v_{i+1} - 2v_i + v_{i-1}}{\Delta^2}, \quad \text{for } i = 2, \ldots I - 1 \tag{24}$$

With the reflecting barrier at $\underline{x}$ and $\bar{x}$,

$$\mathcal{A}v(x_1) = -\frac{\sigma^2}{2\Delta^2}v_1 + \frac{\sigma^2}{2\Delta^2}v_2 \tag{25}$$

$$\mathcal{A}v(x_I) \approx \left(-\frac{\mu_I}{\Delta} + \frac{\sigma_I^2}{2\Delta^2}\right) v_{I-1} + \left(\frac{\mu_I}{\Delta} - \frac{\sigma_I^2}{2\Delta^2}\right) v_I \tag{26}$$

1.3 Stationary Distribution of Reflected Univariate Diffusions

Given a model without any birth/death/stopping, let the pdf of the distribution of x be,

$$\partial_t f(t, x) = \mathcal{A}^* f(t, x), \text{ with appropriate BVs and} \tag{27}$$

$$1 = \int_0^1 f(t, x)\mathrm{d}x, \text{ for all } t \tag{28}$$

where $\mathcal{A}^*$ is the adjoint operator of the $\mathcal{A}$ operator of the stochastic process. In the stationary set, $\partial_t f(t, x) = 0$ for all x. That is,

$$0 = \mathcal{A}^* f(x) \tag{29}$$

$$1 = \int_0^1 f(x)\mathrm{d}x \tag{30}$$

If the linear operator $\mathcal{A}$ is discretized as a matrix A, then it can be shown that the discretized $\mathcal{A}^*$ is the transpose A^T. To find the discretized PDF of the distribution, $f \in \mathbb{R}^I$, solve

$$0 = A^T f \tag{31}$$

$$1 = \omega \cdot f \tag{32}$$

where $\omega \in \mathbb{R}^I$ is a quadrature weight to ensure it integrates to 1. For crude simplicity, could use a vector of 1s. Alternatively, ω could be trapezoidal or simpsons rule weights (depending on if th$x$ grid is uniform).

As an Eigenvalue Problem   Note that (31) can be written as an eigenvalue problem by adding f to both sides,

$$(A^T + \mathbf{I})f = f \tag{33}$$

Therefore, the stationary distribution can be found as the eigenvector of $A^T + \mathbf{I}$ associated with the unity eigenvalue, and rescales to ensure (32) holds. This is useful since A^T is always singular, and hence (31) cannot be solved as a simple linear system.

As a Linear-Least Squares Problem   Alternatively, consider solving (31) and (32) as the following linear least squares problem, subject to $f \geq 0$ inequality constraint.

$$\min_f \|y - Xf\|^2 \tag{34}$$

$$\text{s.t. } f \geq 0 \tag{35}$$

Where,

$$y \equiv \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^{I+1} \tag{36}$$

$$X \equiv \begin{bmatrix} A^T \\ \omega \end{bmatrix} \in \mathbb{R}^{(I+1)\times I} \tag{37}$$

1.4   Solving Simple Time-Independent Linear HJBE

In the simple case of an uncontrolled linear stochastic process with no optimal stopping, if an agent has payoff $u(x)$ with discounting $\rho > 0$, then the following HJBE determines the value

$$\rho v(x) = u(x) + \mathcal{A}v(x) \tag{38}$$

subject to the appropriate boundary conditions. Discretizing the grid, let $u \equiv \{u(x_i)\}_{i=1}^I \in \mathbb{R}^I$, then the discretized version is,

$$\rho v = u + Av \tag{39}$$

Or,

$$(\rho \mathbf{I} - A)v = u \tag{40}$$

This simple linear system can be solved for v. In principle, the HJBE and the KFE could be solved at the same time (here they are uncoupled, but in principle there could be direct connections in a nonlinear setup) by stacking $q \equiv \begin{bmatrix} v & f \end{bmatrix}^T \in \mathbb{R}^{2I}$ (37) and (40) as a linear least squares problem,

$$\min_q \|y - Xq\|^2 \tag{41}$$

Where,

$$y \equiv \begin{bmatrix} u \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^{2I+1} \tag{42}$$

$$X \equiv \begin{bmatrix} \rho\mathbf{I} - A & \mathbf{0}_{I \times I} \\ \mathbf{0}_{I \times I} & A^T \\ \mathbf{0}_{1 \times I} & \omega \end{bmatrix} \in \mathbb{R}^{(2I+1) \times I} \tag{43}$$

If problems occur with negativity of the f, then linear constraints could be added.

2 Time-Dependent Univariate Diffusions

This section implements a version of Section 2, but where important parameters/settings/etc. vary with time. In that case, the SDE is

$$\mathrm{d}x_t = \mu(t, x_t)\mathrm{d}t + \sigma(t, x_t)\mathrm{d}\mathbb{W}_t \tag{44}$$

and the infinitesimal generator is

$$\mathcal{A} \equiv \mu(t, x)\partial_x + \frac{\sigma(t, x)^2}{2}\partial_{xx} + \partial_t \tag{45}$$

Even if the $\mu(\cdot)$ and $\sigma(\cdot)$ are independent of time, the discretized generator used for solving the problem may be dependent on time due to boundary values.

References

KUSHNER, H., AND P. G. DUPUIS (2001): *Numerical methods for stochastic control problems in continuous time*, vol. 24. Springer Science & Business Media, second edition edn.