Derivations and Applications for `SimpleDifferentialOperators.jl`

Chiyoung Ahn and Jesse Perla

April 9, 2019

1 Overview

This package is intended to be a stop-gap while more complete and higher-performance solutions are implemented (in particular, the evolution of `https://github.com/JuliaDiffEq/DiffEqOperators.jl/`). In the meantime, the package hopefully provides a solution for discretizing operators and solving ODEs/PDEs.

The focus is on discretizing linear (and, with some careful checks, affine) operators. In the case of models where the process is non-linear, it is most appropriate for algorithms that iteratively solve linear ODEs/PDEs. Some notation used throughout the document

- Let $x \in R$ is the general "spatial" state variable
- Derivatives are denoted by the operator ∂_x and univariate derivatives such as $\partial_x \tilde{v}(x) \equiv \tilde{v}'(x)$.
- Use the vertical bar to denote operator evaluation at a particular point. That is if $\tilde{B} \equiv \partial_x|_{x=x_0}$ then $\tilde{B}\tilde{v}(x) = \partial_x \tilde{v}(x_0)$, and if $\tilde{B} \equiv 1|_{x=x_0}$ then $\tilde{B}\tilde{v}(x) = \tilde{v}(x_0)$. For the notation, the 1 is simply the identity operator on the function instead of applying a derivative (i.e. $\tilde{B} \equiv 1|_{x=x_0}$ then $\tilde{B}\tilde{v}(\cdot) = 1 \times \tilde{v}(x_0)$)
- Let W_t be the Wiener process with the integral defined by the Ito interpretation

1.1 Linear Differential Equations

ODEs (e.g. Steady State) To understand the class of models that this can support, first look at the sort of ODE that comes out of solving a stationary model. The general pattern is a linear differential operator $\tilde{L}$, a boundary condition operator $\tilde{B}$, the function of interest $\tilde{v}(x)$, and the affine terms $\tilde{f}(\cdot)$ and b. The general problem to solve is to find the $\tilde{v}(x)$ such that.

$$0 = \tilde{L}\tilde{v}(x) - \tilde{f}(x) \tag{1}$$
$$0 = \tilde{B}\tilde{v}(x) - b \tag{2}$$

Motivating Example As a simple example, let

- $x \in [x_{\min}, x_{\max}]$ be a state variable on a domain following the SDE

$$dx_t = \mu dt + \sigma dW_t$$

  where the variable x_t is reflected at $x_{\min}$ and $x_{\max}$

- The payoffs for state x are a function $\tilde{f}(x)$ defines on the domain

- $\tilde{v}(x)$ as the value of the the stream of payoffs discounted at rate $r > 0$

Then, through standard arguments, the stationary Bellman equation along with boundary conditions is

$$r\tilde{v}(x) = \tilde{f}(x) + \mu\partial_x\tilde{v}(x) + \frac{\sigma^2}{2}\partial_{xx}\tilde{v}(x) \tag{3}$$

$$\partial_x\tilde{v}(x_{\min}) = 0 \tag{4}$$

$$\partial_x\tilde{v}(x_{\max}) = 0 \tag{5}$$

Mapping to the notation of (1) and (2)

$$\tilde{L} \equiv r - \mu\partial_x - \frac{\sigma^2}{2}\partial_{xx} \tag{6}$$

$$\tilde{B} \equiv \begin{bmatrix} \partial_x|_{x=x_{\min}} \\ \partial_x|_{x=x_{\max}} \end{bmatrix} \tag{7}$$

$$b \equiv \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{8}$$

This package will allow you to define the $\tilde{L}$, $\tilde{B}$, and b to solve for a discretization of the $\tilde{v}(x)$ function.

PDEs (i.e. Time-Varying) The motivating example above has no time-variation in any of the parameters, payoffs, or boundary conditions. Consider that the operators, payoffs, and boundary conditions could change over time – which we denote with a t subscript. As a variation on (1) and (2)

$$\partial_t\tilde{v}(t,x) = \tilde{L}(t)\tilde{v}(t,x) - \tilde{f}(t,x) \tag{9}$$

$$0 = \tilde{B}(t)\tilde{v}(t,x) - b(t) \tag{10}$$

Subsect to an initial condition, $\tilde{v}(0,x)$ given or potentially a boundary value, $\tilde{v}(T,x)$ for some T.

This is a linear PDE where the operators, boundary conditions, and payoffs all may change over time.

Motivating Example for Dynamics Going back to the motivating example, consider an extension where We will make the following assumptions

- The discount rate, drift, and payoffs could be time varying. i.e. $r(t), \mu(t)$ and $\tilde{f}(t,x)$.
- After some T the system is stationary because $r(t) = r(T), \mu(t) = \mu(T)$ and $\tilde{f}(t,x) = \tilde{f}(T,x)$ for all $t \geq T$

Through standard arguments, the Bellman equation is

$$r\tilde{v}(t,x) = \tilde{f}(t,x) + \mu(t)\partial_x\tilde{v}(x) + \frac{\sigma^2}{2}\partial_{xx}\tilde{v}(x) + \partial_t\tilde{v}(t,x) \tag{11}$$

$$\partial_x\tilde{v}(t,x_{\min}) = 0 \tag{12}$$

$$\partial_x\tilde{v}(t,x_{\max}) = 0 \tag{13}$$

Mapping to the notation of the PDE in (9) and (10)

$$\tilde{L}(t) \equiv r(t) - \mu(t)\partial_x - \frac{\sigma^2}{2}\partial_{xx} \tag{14}$$

$$\tilde{B} \equiv \begin{bmatrix} \partial_x|_{x=x_{\min}} \\ \partial_x|_{x=x_{\max}} \end{bmatrix} \tag{15}$$

$$b \equiv \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{16}$$

In order to find the steady state, we can solve the stationary equation with $\partial_t \tilde{v}(T,x) = 0$ (i.e. with $\tilde{L}(T)$ fixed to find $\tilde{v}(T,x)$) and then use the $\tilde{v}(T,x)$ as a boundary value to solve for the $\tilde{v}(t,x)$ by solving the PDE in

Boundary Conditions The package supports some key boundary conditions used for stochastic processes and ODE/PDEs.

As will become clear in the discretization, whether the boundary condition is homogenous or not (i.e. $b > 0$ or $b = 0$) is important for the numerical methods. If the boundary conditions are inhomogeneous, then the setup is affine. To detail a few of the one-dimensional versions of the supported boundary conditions

- Reflecting Barriers (i.e. a homogeneous Neumann Boundary Conditions)

$$\partial_x \tilde{v}(x_{\min}) = 0 \tag{17}$$
$$\partial_x \tilde{v}(x_{\max}) = 0 \tag{18}$$

  or in operator form

$$\tilde{B} \equiv \begin{bmatrix} \partial_x|_{x=x_{\min}} \\ \partial_x|_{x=x_{\max}} \end{bmatrix} \tag{19}$$
$$b \equiv \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{20}$$

- Mixed boundary conditions (i.e a homogeneous Robin Boundary Conditions):

$$\underline{\xi}\tilde{v}(x_{\min}) + \partial_x \tilde{v}(x_{\min}) = 0 \tag{21}$$
$$\overline{\xi}\tilde{v}(x_{\max}) + \partial_x \tilde{v}(x_{\max}) = 0 \tag{22}$$

  Note that when $\underline{\xi} = \overline{\xi} = 0$, this nests the reflecting barriers. In operator form,

$$\tilde{B} \equiv \begin{bmatrix} \partial_x|_{x=x_{\min}} + \underline{\xi}\, 1|_{x=x_{\min}} \\ \partial_x|_{x=x_{\max}} + \overline{\xi}\, 1|_{x=x_{\max}} \end{bmatrix} \tag{23}$$
$$b \equiv \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{24}$$

- Absorbing Barriers (i.e. a homogenous or inhomogeneous Dirichlet Boundary Conditions)

$$\tilde{v}(x_{\min}) = b_1 \tag{25}$$
$$\tilde{v}(x_{\max}) = b_2 \tag{26}$$

  In the case of $b_1 = b_2 = 0$, this is homogeneous. In operator form,

$$\tilde{B} \equiv \begin{bmatrix} 1|_{x=x_{\min}} \\ 1|_{x=x_{\max}} \end{bmatrix} \tag{27}$$
$$b \equiv \begin{bmatrix} b_0 \\ b_2 \end{bmatrix} \tag{28}$$

Of course, many models would have different boundary conditions on different sides of the domain, which entails mixing and matching rows in the B and b matrices.

2 Discretization

2.1 Notation

This section defines the grids and other notation for the discretization.

- Define an irregular grid $\{x_i\}_{i=0}^{M+1}$ with **boundary nodes**, $x_0 = x_{\min}$ and $x_{M+1} = x_{\max}$. Denote the **extended grid** as $\overline{x} \equiv \{x_i\}_{i=0}^{M+1}$ and the **interior grid**, a collection of nodes excluding the boundary nodes, as $x \equiv \{x_i\}_{i=1}^{M}$.
- Recall that the continuous functions, prior to discretization, are denoted like $\tilde{u}(x)$. The discretization of $\tilde{u}(x)$ on the interior $x \in \mathbb{R}^M$ is denoted $u \equiv \{\tilde{u}(x_i)\}_{i=1}^{M} \in \mathbb{R}^M$ and the discretization of $\tilde{u}(x)$ on $\bar{x} \in \mathbb{R}^{M+2}$ is $\bar{u} \equiv \{\tilde{u}(x_i)\}_{i=0}^{M+1} \in \mathbb{R}^{\bar{M+2}}$.
- When we discretize a particular operator, e.g. $\tilde{L}$, we will drop the tilde to become L. The typical size of this, before applying boundary conditions, is $L \in \mathbb{R}^{M \times (M+2)}$.
- Denote the backward and forward distance between the grid points as

$$\Delta_{i,-} \equiv x_i - x_{i-1}, \text{ for } i = 1, \ldots, M+1 \tag{29}$$

$$\Delta_{i,+} \equiv x_{i+1} - x_i, \text{ for } i = 0, \ldots, M \tag{30}$$

- Define the vector of backwards and forwards first differences, padding with $\Delta_{0,-} = \Delta_{M+1,+} = 0$, as

$$\Delta_- \equiv \begin{bmatrix} 0 \\ \mathrm{diff}(z) \end{bmatrix} \in \mathbb{R}^{M+2} \tag{31}$$

$$\Delta_+ \equiv \begin{bmatrix} \mathrm{diff}(z) \\ 0 \end{bmatrix} \in \mathbb{R}^{M+2} \tag{32}$$

- Some special matrices to help in the composition notation:
  - $\mathbf{I}_N$ is the $N \times N$ identity matrix. Always drop the subscript when the dimensions are unambiguous, as it would be the same in the code
  - $\mathbf{0}_N$ is the column vector of N 0s, and $\mathbf{0}_N^\top$ a row vector
  - $\mathbf{0}_{N \times M}$ is the $N \times M$ matrix of 0s
  - See Appendix A for the definitions of $\mathrm{toep}(\cdot), \mathrm{band}_{l,u}^{n,m}(\cdot)$ and $\mathrm{tridiag}(\cdot)$ matrices

In order to discretize these operators with finite-differences, we need to choose a stencil. Denote the discretization

- For the first-derivative operator $\tilde{L}_1 \equiv \partial_x$, denote the stencils for discretizing with backwards and forward first-differences respectively as $L_{1-} \in \mathbb{R}^{M \times (M+2)}$ and $L_{1+} \in \mathbb{R}^{M \times (M+2)}$.[1]
- For the second-derivative operator $\tilde{L}_2 \equiv \partial_{xx}$, always use central differences and denote the discretized operator as $L_2 \in \mathbb{R}^{M \times (M+2)}$

For first-derivatives, the choice of L_{1-} vs. L_{1+} or a combination of them, will use upwind finite differences.

[1] Note that the stencil for both of these only really needs to be defined on $M \times (M+1)$ but we will pad a column with 0s to make composition easier. In the current form, the package composes operators as sparse matrices. Depending on the circumstances, this code will execute slower than a hand-tweaked model creating composed operators directly. In many cases, this wouldn't be a problem, but in some algorithms where operators need to be redefined frequently in tight loops, it might be. In those cases, use the output of this package for test-suites on hand-built discretizations.

3 Discretizing Operators with a Regular Grid

In this section, we study discretization schemes under regular grids, i.e., grids such that $x_{i+1} - x_i = \Delta$ for all $i = 0, ..., M$ for some fixed $\Delta > 0$.

Throughout, take a function of interest $\tilde{v}(x)$ defined on the grid, and define $\bar{v} \equiv \{\tilde{v}(x_i)\}_{i=0}^{M+1}$ and $v \equiv \{\tilde{v}(x_i)\}_{i=1}^{M}$.

3.1 Discretized Operators

In this section we derive the stencils for operators of various orders.

First Derivative Operators To discretize the $\tilde{L}_1$ operator, we can use a backward difference approximation

$$\tilde{L}_1 \tilde{v}(x_i) \equiv \partial_x \tilde{v}(x_i) \approx \frac{\bar{v}_i - \bar{v}_{i-1}}{\Delta}, \text{ for } i = 1, \dots M \tag{33}$$

And with forward differences

$$\tilde{L}_1 \tilde{v}(x_i) \equiv \partial_x \tilde{v}(x_i) \approx \frac{\bar{v}_{i+1} - \bar{v}_i}{\Delta}, \text{ for } i = 1, \dots M \tag{34}$$

In order to calculate the derivatives for all $i = 1, \dots M$ (i.e. in the interior) we can stack these up and apply to the extension $\bar{v}$.

$$\{\partial_x \tilde{v}(x_i)\}_{i=1}^{M} \approx L_{1-} \cdot \bar{v} \tag{35}$$

or,

$$\{\partial_x \tilde{v}(x_i)\}_{i=1}^{M} \approx L_{1+} \cdot \bar{v} \tag{36}$$

Where we define L_{1-} from applying (33) to the $\bar{v}$ vector for all $i = 1, \dots M$

$$L_{1-} \equiv \frac{1}{\Delta} \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & \ddots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 & 0 \end{bmatrix}_{M \times (M+2)} \tag{37}$$

$$= \text{band}_{0,1}^{M,M+2}(-\mathbf{1}_M, \mathbf{1}_M) \tag{38}$$

And similarly define L_{1+} from applying (34) to the $\bar{v}$ vector for all $i = 1, \dots M$

$$L_{1+} \equiv \frac{1}{\Delta} \begin{bmatrix} 0 & -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix}_{M \times (M+2)} \tag{39}$$

$$= \text{band}_{0,2}^{M,M+2}(\mathbf{0}_M, -\mathbf{1}_M, \mathbf{1}_M) \tag{40}$$

It is important to note that while these operators map the $\bar{v}$ (i.e. including the boundary points), the operator only maps to points on the interior i.e. $i = 1, \dots M$.

Second Derivative Operators To discretize the $\tilde{L}_2$ second order operator, we can use central differences

$$\tilde{L}_2\tilde{v}(x_i) \equiv \partial_{xx}\tilde{v}(x_i) \approx \frac{\bar{v}_{i+1} - 2\bar{v}_i + \bar{v}_{i-1}}{\Delta^2}, \text{ for } i = 1, \dots M \tag{41}$$

In order to calculate the derivatives for all $i = 1, \dots M$ (i.e. in the interior) we can stack these up and apply to the extension $\bar{v}$.

$$\{\partial_{xx}\tilde{v}(x_i)\}_{i=1}^{M} \approx L_2 \cdot \bar{v} \tag{42}$$

Where we define L_2 from applying (41) to the $\bar{v}$ vector for all $i = 1, \dots M$

$$L_2 \equiv \frac{1}{\Delta^2}\begin{bmatrix} 1 & -2 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -2 & 1 \end{bmatrix}_{M\times(M+2)} \tag{43}$$

$$= \text{band}_{0,2}^{M,M+2}(\mathbf{1}_M, -2 \times \mathbf{1}_M, \mathbf{1}_M) \tag{44}$$

Identity Operators For simplicity in composition, also consider the discretization of the identity operator (i.e. not applying any derivatives or stencils). For simplicity, define the identity operator as the 0-th order operator $\tilde{L}_0 \equiv I$ so that $\tilde{L}_0\tilde{v}(x) = \tilde{v}(x)$.

With this, the operator applied to the $\bar{v}$ vector is trivial

$$\tilde{L}_0\tilde{v}(x_i) \equiv \tilde{v}(x_i) = \bar{v}_i, \text{ for } i = 1, \dots M \tag{45}$$

And stacking it up for all $i = 1, \dots M$,

$$\{\tilde{v}(x_i)\}_{i=1}^{M} \approx L_0 \cdot \bar{v} \tag{46}$$

Where

$$L_0 \equiv \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 \end{bmatrix}_{M\times(M+2)} \tag{47}$$

$$= \text{band}_{0,1}^{M,M+2}(\mathbf{0}_M, \mathbf{1}_M) \tag{48}$$

This operator trivially maps from the $\bar{v}$ to extract the interior v and ignoring the boundaries $\bar{v}_0$ and $\bar{v}_{P+1}$. Its primary role will be when composing operators rather than being used directly.

3.2 Representation with Special Matrices

See **??** for more details on the notation of special matrices (e.g. Toeplitz, Tridiagonal, and Banded)

These can be defined in more compact forms as banded matrices

$$L_{1-} \equiv \frac{1}{\Delta}\begin{bmatrix}\text{tridiag}(\mathbf{0}_{M-1}, -\mathbf{1}_M, \mathbf{1}_M) & \mathbf{0}_M\end{bmatrix} \tag{49}$$

$$L_{1+} \equiv \frac{1}{\Delta}\begin{bmatrix}\mathbf{0}_M & \text{tridiag}(\mathbf{0}_{M-1}, -\mathbf{1}_M, \mathbf{1}_M)\end{bmatrix} \tag{50}$$

$$L_2 \equiv \frac{1}{\Delta^2}\text{tridiag}^{+}(\mathbf{1}_M, -2\mathbf{1}_M, \mathbf{1}_M) \tag{51}$$

$$L_0 \equiv \begin{bmatrix}\mathbf{0}_M & \mathbf{I}_M & \mathbf{0}_M\end{bmatrix} \tag{52}$$

where $\text{tridiag}(x, y, z)$ is a matrix whose lower, main, upper diagonal vectors are x, y, and z, respectively, and $\text{tridiag}^{+}(x, y, z)$ is a matrix whose main, upper, and second upper diagonal vectors x, y, and z, respectively.

3.3 Applying Boundary Conditions

Boundary conditions can be applied manually by using operators on extended grids, $\overline{x}$, to find solutions on extended grids. Suppose that we want to solve a system $\tilde{L}\tilde{v}(x) = \tilde{f}(x)$ for $\tilde{v}(x)$ where L is a linear combination of discretized differential operators for some f that represents the values of a function $\tilde{f}(\cdot)$ on discretized interior x. To solve the system under boundary conditions on v, one can construct and solve the following extended system:

$$\begin{bmatrix} L \\ B \end{bmatrix} \overline{v} = \begin{bmatrix} f \\ b \end{bmatrix} \tag{53}$$

with M_E by $(M+2)$ matrix B and M_E-length vector b that represent the current boundary conditions, where M_E is the number of boundary conditions to be applied. The solution of (53), $\tilde{v}(\overline{x})$ can be decomposed into

$$\overline{v} = \begin{bmatrix} \tilde{v}(x_0) \\ v \\ \tilde{v}(x_{M+1}) \end{bmatrix} \tag{54}$$

which also gives the solution for v.

3.3.1 Mixed Boundary Conditions

Recall mixed boundary conditions from (21) and (22). Note that reflecting barrier conditions are special cases with $\overline{\xi} = \underline{\xi} = 0$. Using forward difference and backward difference discretization scheme for the lower bound and upper bound respectively, we have

$$\frac{\overline{v}_1 - \overline{v}_0}{\Delta} + \underline{\xi}\overline{v}_0 = 0 \tag{55}$$

$$\frac{\overline{v}_{M+1} - \overline{v}_M}{\Delta} + \overline{\xi}\overline{v}_{M+1} = 0 \tag{56}$$

Thus, the corresponding boundary condition matrix B is

$$B = \begin{bmatrix} -\frac{1}{\Delta} + \underline{\xi} & \frac{1}{\Delta} & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & -\frac{1}{\Delta} & \frac{1}{\Delta} + \overline{\xi} \end{bmatrix}_{2 \times (M+2)} \qquad b = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{57}$$

which provides the identical system as

$$B = \begin{bmatrix} -1 + \underline{\xi}\Delta & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 + \overline{\xi}\Delta \end{bmatrix}_{2 \times (M+2)} \qquad b = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{58}$$

Adding the first row of $\overline{L}_{1-}$ by the first row of B in (86) multiplied by $(-1 + \underline{\xi}\Delta)^{-1}\Delta^{-1}$ gives, with the corresponding row operation matrix R,

$$RL_{1-} = \frac{1}{\Delta} \begin{bmatrix} 0 & 1 + (-1 + \underline{\xi}\Delta)^{-1} & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & \ddots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 1 & 0 \end{bmatrix}_{M \times (M+2)} \tag{59}$$

note that there is no zero element in the first and last column for nodes on boundaries. Hence, solving the corresponding extended system,

$$\begin{bmatrix} L \\ B \end{bmatrix} = \begin{bmatrix} f \\ b \end{bmatrix} \tag{60}$$

is identical as solving the following system

$$R \begin{bmatrix} L \\ B \end{bmatrix} = R \begin{bmatrix} f \\ b \end{bmatrix} \tag{61}$$

$$= \begin{bmatrix} f \\ b \end{bmatrix} \tag{62}$$

as b is a zero vector so that the row operations R do not change anything on the RHS. Furthermore, limited to the interior, solving v in the system above is identical as solving the following system with an operator L^B on interior nodes:

$$L^B v = f \tag{63}$$

where we have $L = L_{1-}$ and $L^B = L^B_{1-}$ with

$$L^B_{1-} \equiv \frac{1}{\Delta} \begin{bmatrix} 1 + (-1 + \underline{\xi}\Delta)^{-1} & 0 & 0 & \dots & 0 & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix}_{M \times M} \tag{64}$$

instead of solving the full system with boundary conditions. Similarly, subtracting the first row of $\overline{L}_{1+}$ by the second row of B in (86) multiplied by $(1 + \overline{\xi}\Delta)^{-1}\Delta^{-1}$ gives the following differential operator with the boundary condition B applied:

$$L^B_{1+} \equiv \frac{1}{\Delta} \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & -1 + (1 + \overline{\xi}\Delta)^{-1} \end{bmatrix}_{M \times M} \tag{65}$$

And by substracting the first row of L_2 by the first row of B multiplied by $(-1 + \underline{\xi}\Delta)^{-1}$ and the last row of L_2 by the second row of B multiplied by $(1 + \overline{\xi}\Delta)^{-1}$, we have the following differential operator with the boundary condition B applied for L_2:

$$L^B_2 \equiv \frac{1}{\Delta^2} \begin{bmatrix} -2 - (-1 + \underline{\xi}\Delta)^{-1} & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -2 + (1 + \overline{\xi}\Delta)^{-1} \end{bmatrix}_{M \times M} \tag{66}$$

which can be defined in more compact forms:

$$L^B_{1-} \equiv \text{tridiag}(-\mathbf{1}_{M-1}, \begin{bmatrix} 1 + (-1 + \underline{\xi}\Delta)^{-1} & \mathbf{1}^T_{M-1} \end{bmatrix}^T, \mathbf{0}_{M-1}) \tag{67}$$

$$L^B_{1+} \equiv \text{tridiag}(\mathbf{0}_{M-1}, \begin{bmatrix} -\mathbf{1}^T_{M-1} & -1 + (1 + \overline{\xi}\Delta)^{-1} \end{bmatrix}^T, \mathbf{1}_{M-1}) \tag{68}$$

$$L^B_2 \equiv \text{tridiag}(\mathbf{1}_{M-1}, \begin{bmatrix} -2 - (-1 + \underline{\xi}\Delta)^{-1} & -2\mathbf{1}^T_{M-2} & -2 + (1 + \overline{\xi}\Delta)^{-1} \end{bmatrix}^T, \mathbf{1}_{M-1}) \tag{69}$$

3.3.2 Absorbing Boundary Conditions

To apply an absorbing barrier condition $\tilde{v}(x_{\min}) = S$ for some $S \in \mathbb{R}$, with one reflecting barrier condition on the upper bound $v'(x_{\max}) = 0$, one can use

$$B = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}_{2\times(M+2)} \qquad b = \begin{bmatrix} S \\ 0 \end{bmatrix} \tag{70}$$

Similarly, one can apply an absorbing condition on the upper bound $\tilde{v}(x_{\max}) = S$ for some $S \in \mathbb{R}$ and the

$$B = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix}_{2\times(M+2)} \qquad b = \begin{bmatrix} 0 \\ S \end{bmatrix} \tag{71}$$

4 Discretizing Operators with a Irregular Grid

4.1 Applying Boundary Conditions

Instead of solving (53) for a value function $\tilde{v}(\overline{x})$ on the extended grid, one can perform Gaussian elimination to reduce the system and solve $\tilde{v}(x)$, which gives the identical solution as the interior of $\tilde{v}(\overline{x})$.

4.2 Irregular grids

Define the vectors of backward and forward distance for interior nodes as follows:

$$\Delta_-^\circ = \{\Delta_{i,-}\}_{i=1}^M \tag{72}$$

$$\Delta_+^\circ = \{\Delta_{i,+}\}_{i=1}^M \tag{73}$$

We can then define the following operators on $\overline{x}$:

$$L_{1-} \equiv \begin{bmatrix} -\Delta_{1,-}^{-1} & \Delta_{1,-}^{-1} & 0 & \dots & 0 & 0 & 0 \\ 0 & -\Delta_{2,-}^{-1} & \Delta_{2,-}^{-1} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \Delta_{M-1,-}^{-1} & 0 & 0 \\ 0 & 0 & 0 & \cdots & -\Delta_{M,-}^{-1} & \Delta_{M,-}^{-1} & 0 \end{bmatrix}_{M\times(M+2)} \tag{74}$$

$$L_{1+} \equiv \begin{bmatrix} 0 & -\Delta_{1,+}^{-1} & \Delta_{1,+}^{-1} & \dots & 0 & 0 & 0 \\ 0 & 0 & -\Delta_{2,+}^{-1} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\Delta_{M-1,+}^{-1} & \Delta_{M-1,+}^{-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\Delta_{M,+}^{-1} & \Delta_{M,+}^{-1} \end{bmatrix}_{M\times(M+2)} \tag{75}$$

$$L_2 \equiv 2 \begin{pmatrix} (\Delta_{1,+}+\Delta_{1,-})^{-1}\Delta_{1,-}^{-1} & -\Delta_{1,-}^{-1}\Delta_{1,+}^{-1} & (\Delta_{1,+}+\Delta_{1,-})^{-1}\Delta_{i,+}^{-1} & \dots & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & (\Delta_{M,+}+\Delta_{M,-})^{-1}\Delta_{M,-}^{-1} & -\Delta_{M,-}^{-1}\Delta_{M,+}^{-1} & (\Delta_{M,+}+\Delta_{M,-})^{-1}\Delta_{M,+}^{-1} \end{pmatrix}_{M\times(M+2)} \tag{76}$$

Note that we use the following discretization scheme scheme from **?**:

$$v''(x_i) \approx \frac{\Delta_{i,-}\tilde{v}(x_i+\Delta_{i,+}) - (\Delta_{i,+}+\Delta_{i,-})\tilde{v}(x_i) + \Delta_{i,+}\tilde{v}(x_i-\Delta_{i,-})}{\frac{1}{2}(\Delta_{i,+}+\Delta_{i,-})\Delta_{i,+}\Delta_{i,-}}, \text{for } i = 1,\dots,M \tag{77}$$

for second-order derivatives.

And one for identity matrix:

$$I \equiv \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}_{M\times(M+2)} \tag{78}$$

Alternatively, in a more compact form, using vectors distances for interior nodes:

$$L_{1-} \equiv \left[\text{tridiag}(\mathbf{0}_{M-1},\, -(\Delta_-^\circ)^{-1},\, (\Delta_-^\circ)^{-1}) \quad \mathbf{0}_M\right] \tag{79}$$

$$L_{1+} \equiv \left[\mathbf{0}_M \quad \text{tridiag}(\mathbf{0}_{M-1},\, -(\Delta_-^\circ)^{-1},\, (\Delta_-^\circ)^{-1})\ \right] \tag{80}$$

$$L_2 \equiv 2 \odot \text{tridiag}^+ \left[(\Delta_-^\circ + \Delta_+^\circ)^{-1} \odot (\Delta_-^\circ)^{-1}, -(\Delta_-^\circ \odot \Delta_+^\circ)^{-1}, (\Delta_-^\circ + \Delta_+^\circ)^{-1} \odot (\Delta_+^\circ)^{-1}\right] \tag{81}$$

$$I \equiv \left[\mathbf{0}_M \quad \text{diag}(\mathbf{1}_M) \quad \mathbf{0}_M\right] \tag{82}$$

4.2.1 Mixed boundary conditions

Recall mixed boundary conditions from (21) and (22). Note that reflecting barrier conditions are special cases with $\overline{\xi} = \underline{\xi} = 0$. Using forward difference and backward difference discretization scheme for the lower bound and upper bound respectively, we have

$$\frac{\overline{v}_1 - \overline{v}_0}{\Delta_{0,+}} - \underline{\xi}\overline{v}_0 = 0 \tag{83}$$

$$\frac{\overline{v}_{M+1} - \overline{v}_M}{\Delta_{M+1,-}} - \overline{\xi}\overline{v}_{M+1} = 0 \tag{84}$$

Thus, the corresponding boundary condition matrix B is

$$B = \begin{bmatrix} -\dfrac{1}{\Delta_{0,+}} + \underline{\xi} & \dfrac{1}{\Delta_{0,+}} & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & -\dfrac{1}{\Delta_{M+1,-}} & \dfrac{1}{\Delta_{M+1,-}} + \overline{\xi} \end{bmatrix}_{2\times(M+2)} \qquad b = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{85}$$

which provides the identical system as

$$B = \begin{bmatrix} -1 + \underline{\xi}\Delta_{1,-} & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 + \overline{\xi}\Delta_{M,+} \end{bmatrix}_{2\times(M+2)} \qquad b = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{86}$$

since $\Delta_{0,+} = \Delta_{1,-}$ and $\Delta_{M+1,-} = \Delta_{M,+}$

The first columns of all the extension operators above, $\overline{L}_{1-}, \overline{L}_{1+}, \overline{L}_2, \overline{I}$, have non-zero element only in the first rows. Thus, a single Gaussian elimination for the first extension grid will suffice to remove the extended. Likewise, in the last columns of all the extension operators have non-zero element only in the last row.

Adding the first row of $\overline{L}_{1-}$ by the first row of B in (86) multiplied by $(-1+\underline{\xi}\Delta_{1,-})^{-1}\Delta_{1,-}^{-1}$ gives, with the corresponding row operation matrix R for Gaussian elimination,

$$RL_{1-} = \begin{bmatrix} 0 & \Delta_{1,-}^{-1}\left[1+(-1+\underline{\xi}\Delta_{1,-})^{-1}\right] & 0 & \dots & 0 & 0 & 0 \\ 0 & -\Delta_{2,-}^{-1} & \Delta_{2,-}^{-1} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \Delta_{M-1,-}^{-1} & 0 & 0 \\ 0 & 0 & 0 & \dots & -\Delta_{M,-}^{-1} & \Delta_{M,-}^{-1} & 0 \end{bmatrix}_{M\times(M+2)} \tag{87}$$

note that there is no zero element in the first and last column for nodes on boundaries. Hence, solving the corresponding extended system,

$$\begin{bmatrix} L \\ B \end{bmatrix} = \begin{bmatrix} f \\ b \end{bmatrix} \tag{88}$$

is identical as solving the following system

$$R\begin{bmatrix} L \\ B \end{bmatrix} = R\begin{bmatrix} f \\ b \end{bmatrix} \tag{89}$$

$$= \begin{bmatrix} f \\ b \end{bmatrix} \tag{90}$$

as b is a zero vector so that the row operations R do not change anything on the RHS. Furthermore, limited to the interior, solving v in the system above is identical as solving the following system with an operator L^B on interior nodes:

$$L^B v = f \tag{91}$$

where we have $L = L_{1-}$ and $L^B = L_{1-}^B$ with

$$L_{1-}^B \equiv \begin{bmatrix} \Delta_{1,-}^{-1}[1+(-1+\underline{\xi}\Delta_{1,-})]^{-1} & 0 & 0 & \dots & 0 & 0 & 0 \\ -\Delta_{2,-}^{-1} & \Delta_{2,-}^{-1} & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\Delta_{M-1,-}^{-1} & \Delta_{M-1,-}^{-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\Delta_{M,-}^{-1} & \Delta_{M,-}^{-1} \end{bmatrix}_{M\times M} \tag{92}$$

instead of solving the full system with boundary conditions. Similarly, subtracting the first row of $\overline{L}_{1+}$ by the second row of B in (86) multiplied by $(1+\overline{\xi}\Delta_{M,+})^{-1}\Delta_{M,+}^{-1}$ gives the following differential operator with the boundary condition B applied:

$$L_{1+}^B \equiv \begin{bmatrix} -\Delta_{1,+}^{-1} & \Delta_{1,+}^{-1} & 0 & \dots & 0 & 0 & 0 \\ 0 & -\Delta_{2,+}^{-1} & \Delta_{2,+}^{-1} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -\Delta_{M-1,+}^{-1} & \Delta_{M-1,+}^{-1} \\ 0 & 0 & 0 & \dots & 0 & 0 & \Delta_{M,+}^{-1}[-1+(1+\overline{\xi}\Delta_{M,+})^{-1}] \end{bmatrix}_{M\times M} \tag{93}$$

And by substracting the first row of L_2 by the first row of B multiplied by $2(-1+\underline{\xi}\Delta_{1,-})^{-1}(\Delta_{1,+}+\Delta_{1,-})^{-1}\Delta_{1,-}^{-1}$ and the last row of L_2 by the second row of B multiplied by $2(1+\overline{\xi}\Delta_{M,+})^{-1}(\Delta_{M,+}+$

$\Delta_{M,-})^{-1}\Delta_{M,+}^{-1}$, we have the following differential operator with the boundary condition B applied for L_2:

$$L_2^B \equiv 2 \begin{pmatrix} \Xi_1 & (\Delta_{1,+}+\Delta_{1,-})^{-1}\Delta_{1,+}^{-1} & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & (\Delta_{i,+}+\Delta_{i,-})^{-1}\Delta_{i,-}^{-1} & -\Delta_{i,-}^{-1}\Delta_{i,+}^{-1} & (\Delta_{i,+}+\Delta_{i,-})^{-1}\Delta_{i,+}^{-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & (\Delta_{M,+}+\Delta_{M,-})^{-1}\Delta_{M,-}^{-1} & \Xi_M \end{pmatrix}_{M\times M} \tag{94}$$

$$I^B \equiv \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}_{M\times M} \tag{95}$$

where

$$\Xi_1 = -\left[\Delta_{1,-}^{-1}\Delta_{1,+}^{-1} + (-1+\underline{\xi}\Delta_{1,-})^{-1}(\Delta_{1,+}+\Delta_{1,-})^{-1}\Delta_{1,-}^{-1}\right] \tag{96}$$

$$\Xi_M = -\left[\Delta_{M,-}^{-1}\Delta_{M,+}^{-1} - (1+\overline{\xi}\Delta_{M,+})^{-1}(\Delta_{M,+}+\Delta_{M,-})^{-1}\Delta_{M,+}^{-1}\right] \tag{97}$$

Or, alternatively, with vectorized differences:

$$L_{1-}^B \equiv \text{tridiag}\left[-(\Delta_-^\circ)^{-1}[2:M], \left[\Delta_{1,-}^{-1}[1+(-1+\underline{\xi}\Delta_{1,-})^{-1}]; (\Delta_-^\circ)^{-1}[2:M]\right], \mathbf{0}_{M-1}\right] \tag{98}$$

$$L_{1+}^B \equiv \text{tridiag}\left[\mathbf{0}_{M-1}, \left[-(\Delta_-^\circ)^{-1}[1:M-1]; \Delta_{M,+}^{-1}[-1+(1+\overline{\xi}\Delta_{M,+})^{-1}]\right], (\Delta_-^\circ)^{-1}[2:M]\right] \tag{99}$$

$$L_2^B \equiv 2 \odot \text{tridiag}\Big[(\Delta_-^\circ + \Delta_+^\circ)^{-1} \odot (\Delta_-^\circ)^{-1}, \tag{100}$$

$$\left[\Xi_1; -(\Delta_-^\circ \odot \Delta_+^\circ)^{-1}[2:M-1]; \Xi_M\right], \tag{101}$$

$$(\Delta_-^\circ + \Delta_+^\circ)^{-1} \odot (\Delta_+^\circ)^{-1}\Big] \tag{102}$$

$$I^B \equiv \text{diag}(\mathbf{1}_M) \tag{103}$$

4.2.2 Absorbing barrier conditions

To apply an absorbing barrier condition $\tilde{v}(x_{\min}) = S$ for some $S \in \mathbb{R}$, with one reflecting barrier condition on the upper bound $v'(x_{\max}) = 0$, one can use

$$B = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}_{2\times(M+2)} \qquad b = \begin{bmatrix} S \\ 0 \end{bmatrix} \tag{104}$$

Similarly, one can apply an absorbing condition on the upper bound $\tilde{v}(x_{\max}) = S$ for some $S \in \mathbb{R}$ and the

$$B = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix}_{2\times(M+2)} \qquad b = \begin{bmatrix} 0 \\ S \end{bmatrix} \tag{105}$$

Note that the corresponding B and b are identical with regular grid cases.

4.3 Examples

Examples 4.1. Consider $L = L_2$ to solve $Lv = f$ with $M = 3$ under uniform grids $\overline{x} = \{x_0, x_1, x_2, x_3, x_4\}$ and $\Delta = 1$, whose corresponding interior grid is $x = \{x_1, x_2, x_3\}$. This gives

$$L^B = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix} \tag{106}$$

so $L^B v = f$ on the grid x results in the following system

$$\begin{bmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \tilde{v}(x_1) \\ \tilde{v}(x_2) \\ \tilde{v}(x_3) \end{bmatrix} = \begin{bmatrix} \tilde{f}(x_1) \\ \tilde{f}(x_2) \\ \tilde{f}(x_3) \end{bmatrix} \tag{107}$$

For the extended system we have

$$L = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \end{bmatrix} \tag{108}$$

$$B = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \tag{109}$$

Constructing the stacked extended system (53) gives

$$\begin{bmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{v}(x_0) \\ \tilde{v}(x_1) \\ \tilde{v}(x_2) \\ \tilde{v}(x_3) \\ \tilde{v}(x_4) \end{bmatrix} = \begin{bmatrix} \tilde{f}(x_1) \\ \tilde{f}(x_2) \\ \tilde{f}(x_3) \\ 0 \\ 0 \end{bmatrix} \tag{110}$$

Note that substracting the first row of L by (-1) times the first row of B returns an identical system as (110). Likewise, substracting the last row of L by (-1) times the last row of B returns the identical system. Performing the two Gaussian elimination yields the following system:

$$\begin{bmatrix} 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{v}(x_0) \\ \tilde{v}(x_1) \\ \tilde{v}(x_2) \\ \tilde{v}(x_3) \\ \tilde{v}(x_4) \end{bmatrix} = \begin{bmatrix} \tilde{f}(x_2) \\ \tilde{f}(x_3) \\ \tilde{f}(x_4) \\ 0 \\ 0 \end{bmatrix} \tag{111}$$

Note that now we have the first three rows of the coefficient matrix with zero columns on the extended nodes, $\tilde{v}(x_1 - \Delta)$ and $\tilde{v}(x_3 + \Delta)$. Extracting the system corresponding to the first three rows returns the following system, which solves the interior of $\bar{v}$, i.e., v:

$$\begin{bmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \tilde{v}(x_1) \\ \tilde{v}(x_2) \\ \tilde{v}(x_3) \end{bmatrix} = \begin{bmatrix} \tilde{f}(x_1) \\ \tilde{f}(x_2) \\ \tilde{f}(x_3) \end{bmatrix} \tag{112}$$

which is identical as (107).

Examples 4.2. Consider solving (107), but this time with an absorbing barrier condition on the lower bound, $\tilde{v}(x_{\min}) = S$ with a boundary condition matrix

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \tag{113}$$

The corresponding extended system is

$$\begin{bmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{v}(x_0) \\ \tilde{v}(x_1) \\ \tilde{v}(x_2) \\ \tilde{v}(x_3) \\ \tilde{v}(x_4) \end{bmatrix} = \begin{bmatrix} \tilde{f}(x_1) \\ \tilde{f}(x_2) \\ \tilde{f}(x_3) \\ S \\ 0 \end{bmatrix} \tag{114}$$

Note that substracting the first row of L by (-1) times the first row of B returns an identical system as (114). Likewise, substracting the last row of L by (-1) times the last row of B returns the identical system. Performing the two Gaussian elimination yields the following system:

$$\begin{bmatrix} 0 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{v}(x_0) \\ \tilde{v}(x_1) \\ \tilde{v}(x_2) \\ \tilde{v}(x_3) \\ \tilde{v}(x_4) \end{bmatrix} = \begin{bmatrix} \tilde{f}(x_1) - S \\ \tilde{f}(x_2) \\ \tilde{f}(x_3) \\ 0 \\ 0 \end{bmatrix} \tag{115}$$

Extracting the system corresponding to the first three rows returns the following system, which solves the interior of $\bar{v}$, i.e., v:

$$\begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \tilde{v}(x_1) \\ \tilde{v}(x_2) \\ \tilde{v}(x_3) \end{bmatrix} = \begin{bmatrix} \tilde{f}(x_1) - S \\ \tilde{f}(x_2) \\ \tilde{f}(x_3) \end{bmatrix} \tag{116}$$

A Special Matrices and Notation

In the code and algebra, a number of the matrices for discretization have special structure.

Toeplitz Matrices First, a Toeplitz matrix is one in which every descending diagonal is constant. See `https://en.wikipedia.org/wiki/Toeplitz_matrix` for more examples and theory. To create, you need to specify the lower off-diagonals (ordered by rows) and the upper off-diagonals (ordered by columns). They are square only if the number of rows and columns passed in is identical.[2] As an example, to construct

$$vr = \begin{bmatrix} a & f & g & h \end{bmatrix} \in \mathbb{R}^4 \tag{117}$$

$$vc = \begin{bmatrix} a & b & c & d & e \end{bmatrix} \in \mathbb{R}^5 \tag{118}$$

$$\text{toep}(vr, vc) = \begin{bmatrix} a & b & c & d & e \\ f & a & b & c & d \\ g & f & a & b & c \\ h & g & f & a & b \end{bmatrix} \in \mathbb{R}^{4\times 5} \tag{119}$$

Banded Matrices Next, consider a banded matrix (`https://en.wikipedia.org/wiki/Band_matrix`). This is a sparse matrix where only a “small” number of diagonals below and/or above the main diagonal are non-zero. Unlike a Toeplitz matrix, the diagonals need not be constant.[3]

To denote a banded matrix, one needs to specify (1) how many lower and upper diagonals are in the matrix; and (2) the values of those diagonals. For example, a diagonal matrix has 0 lower and 0 upper diagonals while a tridiagonal matrix has 1 of each.

As an example of the notation, first define vectors of diagonals

$$B_{-1} \equiv \begin{bmatrix} B_{21} & B_{32} & B_{43} & B_{54} \end{bmatrix} \in \mathbb{R}^4 \tag{120}$$

$$B_{0} \equiv \begin{bmatrix} B_{11} & B_{22} & B_{33} & B_{44} & B_{55} \end{bmatrix} \in \mathbb{R}^5 \tag{121}$$

$$B_{+1} \equiv \begin{bmatrix} B_{12} & B_{23} & B_{34} & B_{45} & B_{56} \end{bmatrix} \in \mathbb{R}^5 \tag{122}$$

$$B_{+2} \equiv \begin{bmatrix} B_{13} & B_{24} & B_{35} & B_{46} \end{bmatrix} \in \mathbb{R}^4 \tag{123}$$

The $\text{band}_{\ell,u}^{n,m}(\cdot)$ function is used to define a banded matrix given diagonals where the ℓ is number of bands below the diagonal and u is the number of bands above the diagonal. The function then takes a list of the diagonals in order (i.e. lower ones, diagonal bands, then upper band)

$$\text{band}_{1,1}^{5,6}(B_{-1}, B_0, B_{+1}) = \begin{bmatrix} B_{11} & B_{12} & 0 & \cdots & \cdots & 0 \\ B_{21} & B_{22} & B_{23} & \ddots & \ddots & \vdots \\ 0 & B_{32} & B_{33} & B_{34} & \ddots & \vdots \\ \vdots & \ddots & B_{43} & B_{44} & B_{45} & 0 \\ \vdots & \ddots & \ddots & B_{54} & B_{55} & B_{56} \end{bmatrix} \in \mathbb{R}^{5\times 6} \tag{124}$$

[2]This notation is intended to match Matlab’s and Julia’s notation. See `https://www.mathworks.com/help/matlab/ref/toeplitz.html` and `https://github.com/JuliaMatrices/ToeplitzMatrices.jl`. Note that with this interface design, if $vr_1 \neq vr_1$ there is an error.

[3]In principle there could be banded Toeplitz matrices (i.e., only a certain bandwidth of off diagonals are non-zero), but at this point we are unaware of software packages along those lines.

And other with a variation on the columns with no sub-diagonals and two super-diagonals

$$\mathrm{band}^{5,6}_{0,2}(B_0, B_{+1}, B_{+2}) = \begin{bmatrix} B_{11} & B_{12} & B_{13} & 0 & \cdots & 0 \\ 0 & B_{22} & B_{23} & B_{24} & 0 & \vdots \\ \vdots & 0 & B_{33} & B_{34} & B_{35} & 0 \\ \vdots & \ddots & 0 & B_{44} & B_{45} & B_{46} \\ \vdots & \ddots & \ddots & 0 & B_{55} & B_{56} \end{bmatrix} \in \mathbb{R}^{5\times 6} \tag{125}$$

The reason that the (n, m) and (l, u) are both needed is the possibility of additional zeros. For example, simply adding another column

$$\mathrm{band}^{5,7}_{0,2}(B_0, B_{+1}, B_{+2}) = \begin{bmatrix} B_{11} & B_{12} & B_{13} & 0 & \cdots & 0 & 0 \\ 0 & B_{22} & B_{23} & B_{24} & 0 & \vdots & \vdots \\ \vdots & 0 & B_{33} & B_{34} & B_{35} & 0 & 0 \\ \vdots & \ddots & 0 & B_{44} & B_{45} & B_{46} & 0 \\ \vdots & \ddots & \ddots & 0 & B_{55} & B_{56} & 0 \end{bmatrix} \in \mathbb{R}^{5\times 7} \tag{126}$$

Note that a Topelitz matrix where many of the diagonals are 0 can be written as a banded matrix, albeit by dropping the extra structure

$$\mathrm{band}^{4,5}_{1,2}\left(f \times \mathbf{1}_3, a \times \mathbf{1}_4, b \times \mathbf{1}_4, c \times \mathbf{1}_3\right) = \mathrm{toep}\left(\begin{bmatrix} a & f & 0 & 0 \end{bmatrix}, \begin{bmatrix} a & b & c & 0 & 0 \end{bmatrix}\right) \tag{127}$$

$$= \begin{bmatrix} a & b & c & 0 & 0 \\ f & a & b & c & 0 \\ 0 & f & a & b & c \\ 0 & 0 & f & a & b \end{bmatrix} \in \mathbb{R}^{4\times 5} \tag{128}$$

$$\tag{129}$$

Tridiagonal Matrices A final set of matrices are sparse, tridiagonal matrices. This is a particular type of square banded matrix with a single off-diagonal in each direction. As always, the extra structure leads to specialized operations. For example,

$$B_{-1} \equiv \begin{bmatrix} B_{21} & B_{32} & B_{43} & B_{54} \end{bmatrix} \in \mathbb{R}^4 \tag{130}$$

$$B_0 \equiv \begin{bmatrix} B_{11} & B_{22} & B_{33} & B_{44} & B_{55} \end{bmatrix} \in \mathbb{R}^5 \tag{131}$$

$$B_{+1} \equiv \begin{bmatrix} B_{12} & B_{23} & B_{34} & B_{45} \end{bmatrix} \in \mathbb{R}^4 \tag{132}$$

Note in the above that we redefined the B_{+1} vector since otherwise the matrix would not be square. Collecting,

$$\mathrm{tridiag}(B_{-1}, B_0, B_{+1}) = \begin{bmatrix} B_{11} & B_{12} & 0 & \cdots & 0 \\ B_{21} & B_{22} & B_{23} & \ddots & \vdots \\ 0 & B_{32} & B_{33} & B_{34} & \vdots \\ \vdots & \ddots & B_{43} & B_{44} & B_{45} \\ \vdots & \ddots & \ddots & B_{54} & B_{55} \end{bmatrix} \in \mathbb{R}^{5\times 5} \tag{133}$$

Which is a banded square matrix

$$= \text{band}_{1,1}^{5,5}(B_{-1}, B_0, B_{+1}) \tag{134}$$

If the diagonals and off-diagonals are all constant, then it is also a Toeplitz matrix,

$$\text{tridiag}\left(f \times \mathbf{1}_3, a \times \mathbf{1}_4, b \times \mathbf{1}_3\right) = \begin{bmatrix} a & b & 0 & 0 \\ f & a & b & 0 \\ 0 & f & a & b \\ 0 & 0 & f & a \end{bmatrix} \in \mathbb{R}^{4\times 4} \tag{135}$$

$$= \text{toep}(\begin{bmatrix} a & f & 0 & 0 \end{bmatrix}, \begin{bmatrix} a & b & 0 & 0 \end{bmatrix}) \tag{136}$$

$$= \text{band}_{1,1}^{4,4}\left(f \times \mathbf{1}_3, a \times \mathbf{1}_4, b \times \mathbf{1}_3\right) \tag{137}$$

B Derivation by substitution

One can also derive the formula for $L_{1-}^B, L_{1+}^B, L_2^B$ in (64), (65), (66) by substitution. For simplicity, here we focus on the case when we have regular grids, i.e., $x_{i+1} - x_i = \Delta$ for some $\Delta > 0$ for all $i = 0, ..., M$.

Using the backwards first-order difference, (21) implies

$$\frac{\tilde{v}(x_1) - \tilde{v}(x_0)}{\Delta} = -\underline{\xi}\tilde{v}(x_0) \tag{138}$$

i.e.,

$$\tilde{v}(x_0) = \frac{1}{1 - \underline{\xi}\Delta}\tilde{v}(x_1) \tag{139}$$

at the lower bound.

Likewise, (22) under the forwards first-order difference yields

$$\frac{\tilde{v}(x_{M+1}) - \tilde{v}(x_M)}{\Delta} = -\overline{\xi}\tilde{v}(x_{M+1}) \tag{140}$$

i.e.,

$$\tilde{v}(x_{M+1}) = \frac{1}{1 + \overline{\xi}\Delta}\tilde{v}(x_M) \tag{141}$$

at the upper bound.

The discretized central difference of second order under (21) at the lower bound is, by substituting (139) in,

$$\frac{\tilde{v}(x_1 + \Delta) - 2\tilde{v}(x_1) + \tilde{v}(x_{\min})}{\Delta^2} = \frac{\tilde{v}(x_1 + \Delta) - \tilde{v}(x_1)}{\Delta^2} - \frac{1}{\Delta}\frac{\tilde{v}(x_1) - \tilde{v}(x_{\min})}{\Delta} \tag{142}$$

$$= \frac{\tilde{v}(x_1 + \Delta) - \tilde{v}(x_1)}{\Delta^2} + \frac{1}{\Delta}\underline{\xi}\tilde{v}(x_1) \tag{143}$$

$$= \frac{1}{\Delta^2}(-1 + \Delta\underline{\xi})^{-1}\tilde{v}(x_1) + \frac{1}{\Delta^2}\tilde{v}(x_1 + \Delta) \tag{144}$$

Similarly, by (22), we have

$$\frac{\tilde{v}(x_{\max}) - 2\tilde{v}(x_M) + \tilde{v}(x_M - \Delta)}{\Delta^2} = \frac{\tilde{v}(x_M - \Delta) - \tilde{v}(x_M)}{\Delta^2} + \frac{1}{\Delta}\frac{\tilde{v}(x_{\max}) - \tilde{v}(x_M)}{\Delta} \tag{145}$$

$$= \frac{\tilde{v}(x_M - \Delta) - \tilde{v}(x_M)}{\Delta^2} - \frac{1}{\Delta}\overline{\xi}\tilde{v}(x_M) \tag{146}$$

$$= \frac{1}{\Delta^2}(-1 - \Delta\overline{\xi})^{-1}\tilde{v}(x_M) + \frac{1}{\Delta^2}\tilde{v}(x_M - \Delta) \tag{147}$$

at the upper bound.

Thus, the corresponding discretized differential operator L_{1-}, L_{1+}, and L_2 are defined as

$$L_{1-}^B \equiv \frac{1}{\Delta}\begin{bmatrix} 1-(1-\underline{\xi}\Delta)^{-1} & 0 & 0 & \dots & 0 & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix}_{M\times M} \tag{148}$$

$$L_{1+}^B \equiv \frac{1}{\Delta}\begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & -1+(1-\overline{\xi}\Delta)^{-1} \end{bmatrix}_{M\times M} \tag{149}$$

$$L_2^B \equiv \frac{1}{\Delta^2}\begin{bmatrix} -2-(1-\underline{\xi}\Delta)^{-1} & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -2+(1+\overline{\xi}\Delta)^{-1} \end{bmatrix}_{M\times M} \tag{150}$$