Monotone Schemes for Two Factor HJB Equations: Nonzero Correlation

Peter Forsyth[1]

[1]Cheriton School of Computer Science
University of Waterloo

September, 2017
Lisbon

Outline

Need to guarantee numerical scheme converges to viscosity solution

- Sufficient conditions (Barles, Souganidis (1991))
  - Monotone, consistent (in the viscosity sense) and ℓ_∞ stable
- Examples known where seemingly reasonable (non-monotone) discretizations converge to incorrect solution

One stochastic factor, several path dependent factors

- Easy to construct a monotone scheme
  - → Forward-backward differencing, semi-Lagrangian timestepping, policy iteration

But suppose we have two (or more) stochastic factors

→ Not so easy to construct monotone schemes if we have nonzero correlation

Example: two factor uncertain volatility

Suppose we have two stochastic factors S_1, S_2 (equities).

Risk neutral processes:

$$
\begin{aligned}
dS_1 &= rS_1 \; dt + \sigma_1 S_1 \; dW_1, \\
dS_2 &= rS_2 \; dt + \sigma_2 S_2 \; dW_2, \\
r &= \text{ risk free rate} \\
\sigma_i &= \text{ volatility} \\
W_{k=1,2} &= \text{ Wiener processes}
\end{aligned}
\tag{1}
$$

where

$$
\begin{aligned}
d[W_1, W_2] &= \rho \; dt \\
\rho &= \text{ correlation}
\end{aligned}
\tag{2}
$$

HJB PDE

No arbitrage value of a contingent claim $\mathcal{U}(S_1, S_2, \tau = T - t)$

$$\mathcal{U}_\tau = \mathcal{L}(\sigma_1, \sigma_2, \rho)\, \mathcal{U}$$

where

$$\begin{aligned}
&\mathcal{L}(\sigma_1, \sigma_2, \rho)\, \mathcal{U} \\
&\quad = \frac{\sigma_1^2 S_1^2}{2}\, \mathcal{U}_{S_1 S_1} + \frac{\sigma_2^2 S_2^2}{2}\, \mathcal{U}_{S_2 S_2} + r\, \mathcal{U}_{S_1} + r\, \mathcal{U}_{S_2} - r\mathcal{U} \\
&\quad + \underbrace{\rho \sigma_1 \sigma_1 S_1 S_2\, \mathcal{U}_{S_1 S_2}}_{cross\ \ derivative\ \ term}
\end{aligned}$$

And we have the initial condition

$$\mathcal{U}(S_1, S_2, 0) \;\; = \;\; \mathcal{W}(S_1, S_2) = \text{ payoff}$$

Uncertain Volatilities, Correlation

Suppose σ_1, σ_2, ρ are uncertain

Define the set of controls Q

$$Q = \{\sigma_1, \sigma_2, \rho\}$$

With the set of admissible controls $\mathcal{Z}$

$$\begin{aligned} \mathcal{Z} &= [\sigma_{1,\text{min}}, \sigma_{1,\text{max}}] \times [\sigma_{2,\text{min}}, \sigma_{2,\text{max}}] \times [\rho_{\text{min}}, \rho_{\text{max}}] \\ & \sigma_{1,\text{min}} \geq 0, \quad \sigma_{2,\text{min}} \geq 0 \\ & -1 \leq \rho_{\text{min}} \leq \rho_{\text{max}} \leq 1. \end{aligned}$$

Worst case cost of hedging, short, $\mathcal{L}^{\mathcal{Q}} \equiv \mathcal{L}(\sigma_1, \sigma_2, \rho)$

$$\mathcal{U}_\tau = \sup_{Q \in \mathcal{Z}} \mathcal{L}^{\mathcal{Q}} \mathcal{U}$$

Worst case cost of hedging, long

$$\mathcal{U}_\tau = \inf_{Q \in \mathcal{Z}} \mathcal{L}^{\mathcal{Q}} \mathcal{U}$$

Discretization

Localize computational domain

$$(S_1, S_2) \in [0, (S_1)_{\max}] \times [0, (S_2)_{\max}] = \Omega$$

Define a set of nodes, timesteps

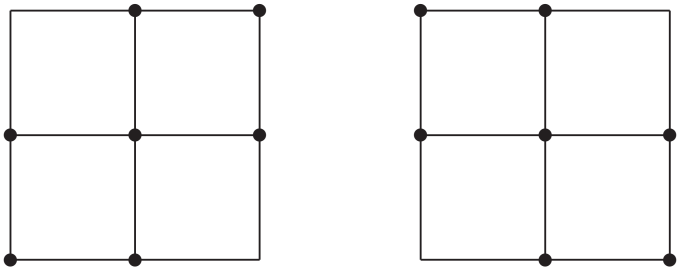
$$\{(S_1)_1, (S_1)_2, \ldots, (S_1)_{N_1}\} \quad ; \quad \{(S_2)_1, (S_2)_2, \ldots, (S_2)_{N_2}\}$$
$$\tau^n = n\Delta\tau, \;\; n = 0, \ldots, N_\tau$$

And a discretization parameter h

$$\max_{(S_1,S_2)\in\Omega} \min_{i,j} \left|(S_1, S_2) - ((S_1)_i, (S_2)_j)\right| = O(h)$$
$$\Delta\tau = O(h)$$

First Attempt: Fixed Stencil

Finite difference of cross-derivative term (seven point stencil)



(a) $\rho \geq 0$ (b) $\rho < 0$

Other terms:

- Three point second derivative finite difference
- Central/forward/backward for first derivative terms
- Try to produce a positive coefficient scheme

Positive Coefficient Scheme

$\mathcal{U}_{i,j}^n \equiv$ approximate solution at $((S_1)_i, (S_2)_j, \tau^n)$

Discretization operator L_f^Q (fixed stencil)[1]

$$
\begin{aligned}
L_f^Q \mathcal{U}_{i,j}^n = & (\alpha_{i,j}^{S_1} - \gamma_{i,j})\mathcal{U}_{i-1,j}^n + (\beta_{i,j}^{S_1} - \gamma_{i,j})\mathcal{U}_{i+1,j}^n \\
& + (\alpha_{i,j}^{S_2} - \gamma_{i,j})\mathcal{U}_{i,j-1}^n + (\beta_{i,j}^{S_2} - \gamma_{i,j})\mathcal{U}_{i,j+1}^n \\
& + \mathbf{1}_{\rho \geq 0}(\gamma_{i,j}\mathcal{U}_{i+1,j+1}^n + \gamma_{i,j}\mathcal{U}_{i-1,j-1}^n) \\
& + \mathbf{1}_{\rho < 0}(\gamma_{i,j}\mathcal{U}_{i+1,j-1}^n + \gamma_{i,j}\mathcal{U}_{i-1,j+1}^n) \\
& - (\alpha_{i,j}^{S_1} + \beta_{i,j}^{S_1} + \alpha_{i,j}^{S_2} + \beta_{i,j}^{S_2} - 2\gamma_{i,j} + r)\mathcal{U}_{i,j}^n
\end{aligned}
$$

Definition 1 (Positive Coefficient Discretization)

L_f^Q is a positive coefficient discretization if $\forall Q \in \mathcal{Z}$

$$
(Red\ \ terms) \geq 0 \quad ; \quad \overbrace{\alpha_{i,j}^{S_k}, \beta_{i,j}^{S_k}, \gamma_{i,j} \geq 0}^{true\ \ by\ \ construction}
$$

[1] Note that α, β, γ are functions of the control Q.

Monotone Schemes

Consider fully implicit timestepping:

$$\mathcal{U}_{i,j}^{n+1} = \mathcal{U}_{i,j}^{n} + \Delta\tau \max_{Q \in Z} \mathcal{L}_f^Q \mathcal{U}_{i,j}^{n+1} \tag{3}$$

which we can write as

$$\mathcal{G}_{i,j}(\mathcal{U}_{i,j}^{n+1}, \mathcal{U}_{i,j}^{n}, \mathcal{U}_{i+1,j}^{n+1}, \ldots) = \mathcal{U}_{i,j}^{n+1} - \mathcal{U}_{i,j}^{n} - \Delta\tau \max_{Q \in Z} \mathcal{L}_f^Q \mathcal{U}_{i,j}^{n+1} = 0 \tag{4}$$

Definition 2 (Monotone Scheme)

Scheme (3) is monotone if $\mathcal{G}_{i,j}(\mathcal{U}_{i,j}^{n+1}, \mathcal{U}_{i,j}^{n}, \mathcal{U}_{i+1,j}^{n+1}, \ldots)$ is a nonincreasing function of neighbours of $\mathcal{U}_{i,j}^{n+1}$, i.e. $(\mathcal{U}_{i,j}^{n}, \mathcal{U}_{i+1,j}^{n+1}, \ldots)$.

Theorem 3 (Positive Coefficient Scheme)

A positive coefficient scheme is monotone.

Conditions for a Positive Coefficient Scheme: Fixed Stencil

Recall that the positive coefficient property has to hold $\forall Q \in \mathcal{Z}$ (i.e. $\alpha_{i,j}, \beta_{i,j}, \gamma_{i,j}$ are functions of Q)

- The problem is the cross-derivative term
- For general $\mathcal{Z}$, this requires severe restrictions on the grid spacing
  - Restricted grid may not allow for fine spacing near strike
- May be impossible to satisfy
- See Reisinger (2016) for a discussion of this.

Alternative: wide stencil method

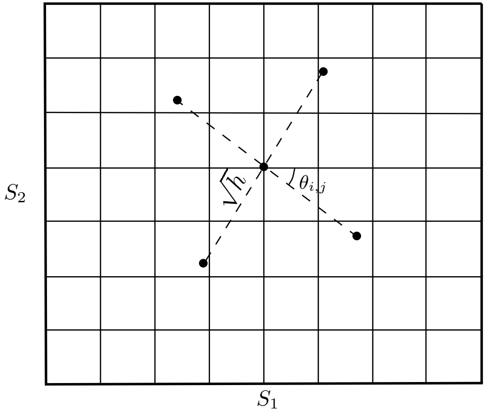
Wide Stencil Method

Wide stencil

- Grid spacing $O(h)$
- At each node, do virtual rotation[2]
  - $\rightarrow$ Choose rotation angle so that local diffusion tensor is diagonal, no cross-derivative term,
  - $\rightarrow$ Finite difference on virtual rotated grid
- Values are interpolated from *real grid*
- Size of virtual stencil $O(\sqrt{h})$
  - We interpolate data for stencil from actual grid
  - Stencil size is $O(\sqrt{h}) \rightarrow$ guarantees consistency

[2] Debrebant and Jakobsen (2013); Reisinger and Rotaetxe Arto (2016); factor the diffusion tensor

Local Rotation



Note: local rotation angle $\theta_{i,j}$ depends on

- Node location, i.e. (S_i, S_j)
- Control Q at this node

Wide Stencil II

Why is this called a wide stencil method?

- Size of (virtual) stencil $O(\sqrt{h})$
- Grid spacing $O(h)$
- Relative stencil length

$$\frac{\sqrt{h}}{h} \to \infty \text{ as } h \to 0$$

What happens near the boundaries?

Simple application of wide stencil

$\rightarrow$ Stencil may require data outside computational domain

Wide Stencil: near boundaries

If we need data $S_1 > (S_1)_{\max}$ or $S_2 > (S_2)_{\max}$

- Localization
  - $\rightarrow$ Use artificial boundary conditions at $(S_1)_{\max}, (S_2)_{\max}$ based on asymptotic form of solution
- Use same asymptotic form for data needed from wide stencil
- Errors small if $(S_1)_{\max}, (S_2)_{\max}$ sufficiently large

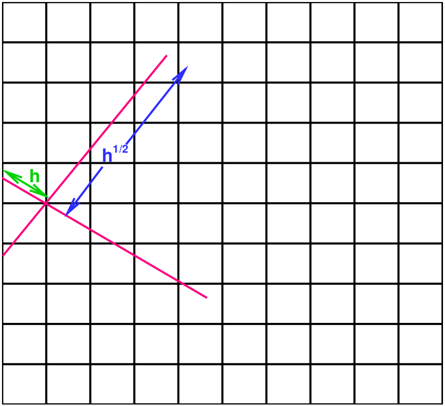
But, what about near $S_1 = 0, S_2 = 0$?

- Wide stencil may need data for $S_1 < 0$ or $S_2 < 0$

Solution:

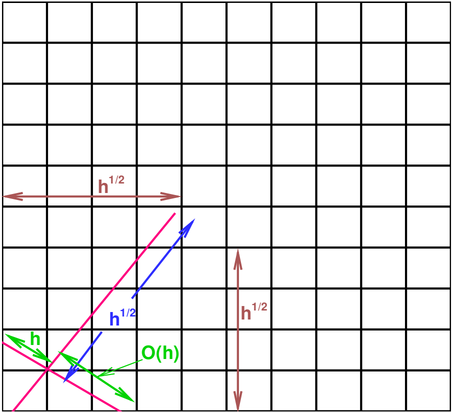
- Shrink stencil arm so that we do not go outside domain

Shrink Stencil Arm



if $(S_1)_i > \sqrt{h}$ or $(S_2)_j > \sqrt{h}) \Rightarrow$ discretization is consistent $O(\sqrt{h})$

What about lower left corner?



Discretization of 2nd order derivative inconsistent here $O(1)$

- Region $(S_1, S_2) \in [0, \sqrt{h}] \times [0, \sqrt{h}]$
- Equation coefficient $O(h) \rightarrow$ consistent discretization of PDE!

Are we just lucky in this case?

If Fichera condition $\rightarrow$ boundary condition required

- We use the necessary specified boundary condition $\rightarrow$ no interpolation error at the truncated stencil points in the rotated stencil $\rightarrow$ consistent

If PDE degenerates in both directions near the *corner*, and no boundary conditions required (this talk)

- Equation coefficient tends to zero $\rightarrow$ consistent

Conjecture 1

Truncating the stencil near the boundary is always consistent.

Proof.

(Maybe)

If no boundary condition required in one direction, but boundary condition required in the other direction, then the virtual local grid *rotates* to align with original grid (as $h \rightarrow 0$) $\Rightarrow$ consistent □

Convergence of wide stencil method $\mathcal{L}_w^Q$

Lemma 4 (Ma and Forsyth (2017))

The fully implicit wide stencil scheme

$$\mathcal{U}_{i,j}^{n+1} = \mathcal{U}_{i,j}^{n} + \Delta\tau \sup_{Q \in \mathcal{Z}} \mathcal{L}_w^Q \mathcal{U}_{i,j}^{n+1}$$

is consistent (in the viscosity sense), ℓ_∞ stable and monotone.

Theorem 5 (Convergence)

The wide stencil method converges to the viscosity solution of the uncertain volatility HJB PDE.

Proof.

The HJB PDE satisfies the strong comparison property (Guyon and Henry-Labordere (2011)). Result follows from Lemma 4 and (Barles and Souganidis (1993)). □

Hybrid Method

Algorithm 1 Hybrid Discretization Method $(\mathcal{L}_H^Q)_{i,j}$

for $i = 1, \ldots, N-1;\ j = 1, \ldots, N_2$ **do**
  if $(\mathcal{L}_f^Q)_{i,j}$ monotone $\forall Q \in \mathcal{Z}$ **then**
    Use fixed stencil at this node $(\mathcal{L}_H^Q)_{i,j} = (\mathcal{L}_f^Q)_{i,j}$
  else
    Use wide stencil at this node $(\mathcal{L}_H^Q)_{i,j} = (\mathcal{L}_w^Q)_{i,j}$
  end if
end for

Fixed stencil used as much as possible (more accurate).

- We do not enforce any grid conditions
- We simply check to see if the monotonicity conditions are satisfied at a given node
- Algorithm 1 only done once at start

Fully Implicit Timestepping

$$\mathcal{U}_{i,j}^{n+1} = \mathcal{U}_{i,j}^{n} + \Delta\tau \sup_{Q \in \mathcal{Z}} \mathcal{L}_H^Q \, \mathcal{U}_{i,j}^{n+1}$$

$$\sup_{Q \in \mathcal{Z}} \left[-(1 - \Delta\tau \mathcal{L}_H^Q) \, \mathcal{U}_{i,j}^{n+1} + \mathcal{U}_{i,j}^{n} \right] = 0$$

Define:

$$\mathbf{U}^n = (\mathcal{U}_{1,1}^n, \mathcal{U}_{2,1}^n, \ldots, \mathcal{U}_{N_1,1}^n, \ldots, \mathcal{U}_{1,N_2}^n, \ldots, \mathcal{U}_{N_1,N_2}^n)$$
$$\mathbf{U}_\ell^n = \mathcal{U}_{i,j}^n, \quad \ell = i + (j-1)N_1.$$

Similarly the vector of optimal controls is

$$\mathcal{Q} = (Q_{1,1}, \ldots, Q_{N_1 N_2})$$

The nonlinear algebraic equations are then[3]

$$\sup_{Q \in \mathcal{Z}} \left\{ -\mathbf{A}(\mathcal{Q})\mathbf{U}^{n+1} + \mathbf{C}(\mathcal{Q}) \right\} = 0, \tag{5}$$

$\mathbf{A}(\mathcal{Q}) =$ matrix of discretized equations ; $\mathbf{C}(\mathcal{Q}) =$ rhs vector

[3] Row ℓ of $\mathbf{A}, \mathbf{C}$ depends only on Q_ℓ

Policy Iteration[4]

Algorithm 2 Policy Iteration

1: Let $(\hat{\mathbf{U}})^0$ = Initial estimate for $\mathbf{U}^{n+1}$
2: **for** $k = 0, 1, 2, \ldots$ until converge **do**
3: $\quad \mathcal{Q}_\ell^k = \arg\max_{\mathcal{Q}_\ell \in \mathcal{Z}} \left\{ -[\mathbf{A}(\mathcal{Q})]\hat{\mathbf{U}}^k + \mathbf{C}(\mathcal{Q}) \right\}_\ell$
4: $\quad$ Solve $[\mathbf{A}(\mathcal{Q}^k)]\hat{\mathbf{U}}^{k+1} = \mathbf{C}(\mathcal{Q}^k)$
5: $\quad$ **if** converged **then**
6: $\quad\quad$ break from the iteration
7: $\quad$ **end if**
8: **end for**

[4]Use ILU-PCG method to solve matrix, complexity $= O((N_1 N_2)^{5/4})$, due to shrunk stencil near boundary and fixed stencil nodes.

Policy Iteration II

Let $(\hat{\mathbf{U}})^0$ = Initial estimate for $\mathbf{U}^{n+1}$
for $k = 0, 1, 2, \ldots$ until converge **do**
$\quad \mathcal{Q}_\ell^k = \underset{\mathcal{Q}_\ell \in \mathcal{Z}}{\arg\max} \left\{ -[\mathbf{A}(\mathcal{Q})]\hat{\mathbf{U}}^k + \mathbf{C}(\mathcal{Q}) \right\}_\ell$
$\quad$ Solve $[\mathbf{A}(\mathcal{Q}^k)]\hat{\mathbf{U}}^{k+1} = \mathbf{C}(\mathcal{Q}^k)$
$\quad$ **if** converged **then**
$\quad\quad$ break from the iteration
$\quad$ **end if**
end for

Theorem 6 (Convergence of Policy Iteration)

If $\forall \mathcal{Q} \in \mathcal{Z}$, $[\mathbf{A}(\mathcal{Q})]$ *is an* $\mathcal{M}$ *matrix, then Policy iteration converges to the unique solution of equation (5).*

For wide stencil nodes

- The rotation angle is a function of $\mathcal{Q}$
  - $\rightarrow$ The stencil changes at each policy iteration

But, we can still prove policy iteration converges!

- Positive coefficient $\rightarrow$ $\mathbf{A}(Q)$ is an $\mathcal{M}$ matrix

Numerical Example (nonconvex payoff)

Butterfly on maximum (worst case short)

$$
\begin{aligned}
S_{\max} &= \max(S_1, S_2), \\
\text{Payoff } &= \max(S_{\max} - K_1, 0) + \max(S_{\max} - K_2, 0) \\
&\quad - 2\max(S_{\max} - (K_1 + K_2)/2, 0).
\end{aligned}
$$

Parameter	Value
Time to expiry (T)	0.25
r	0.05
σ_1	[.3, .5]
σ_2	[.3, .5]
ρ	[.3, .5]
K_1	34
K_1	46

Grid/timesteps

Refine Level	Timesteps	S_1 nodes	S_2 nodes	$\partial\mathcal{Z}$ nodes
1	25	91	91	24
2	50	181	181	46
3	100	361	361	90
4	200	721	721	178

For fixed stencil, analytic expression for global maximum of objective function on $\partial\mathcal{Z}$.

For wide stencil[5], need to discretize control and do linear search[6] on $\partial\mathcal{Z}$.

[5] The policy iteration matrix is a discontinuous function of the control in this case.

[6] The cost of the linear search far exceeds the cost of solving the matrix at each Policy iteration

Convergence study

	Hybrid Scheme		Pure Wide Stencil	
Refine	Value	Diff	Value	Diff
1	2.7160		2.6371	
2	2.6946	0.0214	2.6397	0.0026
3	2.6880	0.0066	2.6650	0.0252
4	2.6862	0.0018	2.6744	0.0094

Table : Butterfly call on max of two, worst case short, value at $(S_1 = S_2 = 40, t = 0)$

Refine	Average policy itns per step		Fraction Fixed
	Hybrid Scheme	Pure Wide	(Hybrid)
1	4.0	3.7	0.38
2	3.8	3.7	0.42
3	3.6	3.6	0.44
4	3.3	3.3	0.45

Table : Fraction fixed is *not small*, consistent with analysis in Reisinger (2016).

Rotation vs. Factoring[7]

Construct virtual local grid with no cross-derivative terms:

- **Rotate** the local grid
- **Factor** the diffusion tensor

Let the diffusion tensor be

$$\mathbf{D} = \frac{1}{2} \begin{pmatrix} \sigma_1^2 S_1^2 & \rho\sigma_1\sigma_2 S_1 S_2 \\ \rho\sigma_1\sigma_2 S_1 S_2 & \sigma_2^2 S_2^2 \end{pmatrix} \tag{6}$$

Factoring **D**

$$\mathbf{D} = \frac{1}{2} \mathbb{C}^T \mathbb{C} \tag{7}$$

Define virtual coordinate system using columns of $\mathbb{C}$

$\rightarrow$ the cross-derivative terms are eliminated

[7] If you are a stochastic process person, factoring is natural, if you are a PDE person, rotation is natural

Rotation vs. Factoring II

Numerical tests:

Refine	Rotation	Factoring
1	2.7160	2.8518
2	2.6946	2.7733
3	2.6880	2.7282
4	2.6862	2.7085

Table : Butterfly call on max of two, worst case short, value at $(S_1 = S_2 = 40, t = 0)$. Hybrid scheme.

- Rotation seems to converge faster than factoring
  - Rotated grid $\rightarrow$ orthogonal
  - Factored grid $\rightarrow$ non-orthogonal

Summary: Uncertain Volatility

- Cross derivative term $\rightarrow$ difficult to construct monotone scheme
- Wide stencil method
  $\rightarrow$ Unconditionally monotone, but only first order
- Hybrid scheme: use fixed stencil as much as possible
  $\rightarrow$ Multi-d generalization of *central differencing as much as possible*
- Empirical results:
  - Local grid rotation better than factoring
  - Hybrid better than pure wide stencil
- Conjecture: truncation of rotated stencil near boundary always consistent

Conclusions

- Wide stencil idea can be easily combined with semi-Lagrangian timestepping if control appears in diffusion and first order terms
  - → See portfolio allocation example (Ma and Forsyth (2016))
- Similar method for multi-factor impulse control
- Implicit discretization, no timestep restriction due to stability
  - Policy iteration rapidly convergent
  - Matrix easy to solve with an iterative method (M-matrix, local orthogonal grid)[8]
- Low accuracy control → accurate value function
- Challenges:
  - Higher dimensions
  - Wide stencil only 1st order
  - Solution of local optimization (need global optimum to $O(h)$)[9]

[8] But see Reisinger and Rotaetxe Arto (2016). Note that we use rotation and $ILU(1)$.

[9] Currently: discretize control, exhaustive search, most costly part of algorithm