Analysis of a Penalty Method for Guaranteed Minimum Withdrawal Benefits (GMWB)

Peter Forsyth[1] Irene Huang[2]

[1]Cheriton School of Computer Science
University of Waterloo

[2]Electrical and Computer Engineering
University of Waterloo

Toronto, March 22-24, 2010

Retirement Risk Zone

Consider an investor with a retirement account, which is invested in the stock market. The investor retires, and makes withdrawals from the retirement account.

The outcomes will be very different in the cases:

- in the first few years after retirement, the market has losses, and the account is further depleted by withdrawals, followed by some years of good market returns; compared to
- a few years of good market returns, after retirement (including withdrawals), followed by some years of losses

Losses in the early years of retirement can be devastating in the long run! Early bad returns can cause complete depletion of the account.

A Typical GMWB Example

Investor pays \$100 to an insurance company, which is invested in a risky asset.

Denote amount in risky asset sub-account by $W = 100$.

The investor also has a virtual guarantee account $A = 100$.

Suppose that the contract runs for 10 years, and the guaranteed withdrawal rate is \$10 per year.

A Typical GMWB Example II

At the end of each year, the investor can choose to withdraw up to \$10 from the account. If $\$\hat{\gamma} \in [0, 10]$ is withdrawn, then

$$
\begin{aligned}
W_{new} &= \max(W_{old} - \hat{\gamma}, 0) \quad ; \quad \textit{Actual investment} \\
A_{new} &= A_{old} - \hat{\gamma} \quad ; \quad \textit{Virtual account}
\end{aligned}
$$

This continues for 10 years. At the end of 10 years, the investor can withdraw anything left, i.e. $\max(W_{new}, A_{new})$.

Note: the investor can continue to withdraw cash as long as $A > 0$, *even if* $W = 0$ (recall that W is invested in a risky asset).

Why is this useful?

The investor can participate in market gains, but still has a guaranteed cash flow, in the case of market losses.

This insulates pensioners from losses in the early years of retirement.

This protection is paid for by deducting a yearly fee η from the amount in the risky account W each year.

The simple form of GMWB described has many variants in practice: Guaranteed Lifetime Withdrawal Benefit (GLWB), ratchet increase of virtual account A if no withdrawals, etc.

We will keep things simple here, and look at the basic GMWB.

Most variable annuities sold in North America have some type of market guarantee.

Some More Details

The investor can choose to withdraw up to the specified contract *rate* G_r without penalty.

Usually, a penalty $(\kappa > 0)$ is charged for withdrawals above G_r.

Let γ be the rate of withdrawal selected by the holder.

Then, the rate of actual cash received by the holder of the GMWB is

$$f(\gamma) = \begin{cases} \gamma & \text{if } 0 \leq \gamma \leq G_r, \\ (1-\kappa)\gamma + \kappa G_r & \text{if } \gamma > G_r. \end{cases}$$

Stochastic Process

Let S denote the value of the risky asset, we assume that the risk neutral process followed by S is

$$\begin{aligned} dS &= rSdt + \sigma SdZ \\ & \quad r = \text{ risk free rate;} \quad \sigma = \text{ volatility} \\ & \quad dZ = \phi\sqrt{dt} \;\; ; \;\; \phi \sim \mathcal{N}(0,1) \end{aligned}$$

The risk neutral process followed by W is then (including withdrawals dA).

$$\begin{aligned} dW &= (r-\eta)Wdt + \sigma WdZ + dA, \quad \text{if } W > 0 \\ dW &= 0, \quad \text{if } W = 0 \\ & \quad \eta = \text{ fee paid for guarantee} \;\; ; \;\; A = \text{ guarantee account} \end{aligned}$$

No-arbitrage Value

Let $V(W, A, \tau)$ ($\tau = T - t$, T is contract expiry) be the no-arbitrage value of the GMWB contract (i.e. the cost of hedging).

At contract expiry ($\tau = 0$) we have (payoff = withdrawal)

$$V(W, A, \tau = 0) \quad = \quad \max(W, A(1 - \kappa))$$

It turns out that it is optimal to withdraw at a rate γ

- $\gamma \in [0, G_r]$, or
- $\gamma = \infty$ (instantaneously withdraw a finite amount)

Singular Control

Let

$$\begin{aligned} \mathcal{L}V &= \frac{1}{2}\sigma^2 W^2 V_{WW} + (r-\eta)WV_W - rV \\ \mathcal{F}V &= 1 - V_W - V_A \end{aligned}$$

Then, as shown in (Dai et al (2008)), the no arbitrage value of this guarantee is given from the solution to the HJB VI

$$\min\left[V_\tau - \mathcal{L}V - G_r \max(\mathcal{F}V, 0), \kappa - \mathcal{F}V\right] = 0 \qquad (1)$$

An Informal Derivation

One way to derive this HJB VI is to consider the optimal control problem with a large, but finite, maximum withdrawal rate $1/\varepsilon$. Let

$$h(\gamma) = \begin{cases} \gamma - \gamma(V_W + V_A) & \text{if } 0 \le \gamma \le G_r, \\ (1-\kappa)\gamma + \kappa G_r - \gamma(V_W + V_A) & \text{if } \gamma > G_r. \end{cases} \tag{2}$$

Then the control problem (for bounded rate $1/\varepsilon$) is

$$V_\tau = \mathcal{L}V + \max_{\gamma \in [0, 1/\varepsilon]} h(\gamma), \tag{3}$$

Noting that the maximum in (3) occurs at $\gamma = \{0, G_r, 1/\varepsilon\}$, a bit of algebra shows that this equation becomes

$$-V_\tau + \mathcal{L}V + \max\left[G_r \max(0, \mathcal{F}V), \frac{(\mathcal{F}V - \kappa)}{\varepsilon} + \kappa G_r \right] = 0\ . \tag{4}$$

HJB Variational Inequality

A more compact form of this equation is

$$\min\left[V_\tau - \mathcal{L}V - G_r \max(0, \mathcal{F}V), V_\tau - \mathcal{L}V - \kappa G_r + \frac{(\kappa - \mathcal{F}V)}{\varepsilon}\right] = 0 \ .$$

Or, since $\varepsilon > 0$,

$$\min\left[V_\tau - \mathcal{L}V - G_r \max(0, \mathcal{F}V), \kappa - \mathcal{F}V + \varepsilon\left(V_\tau - \mathcal{L}V - \kappa G_r\right)\right] = 0 \ .$$

Let $\varepsilon \to 0$

$$\min\left[V_\tau - \mathcal{L}V - G_r \max(\mathcal{F}V, 0), \kappa - \mathcal{F}V\right] = 0 \ . \qquad (5)$$

which is the final form of the HJB equation.

Basic Idea: Penalty Method

Keeping ε finite, we can rewrite (4) in *control form*, in terms of the controls $\{\varphi, \psi\}$

$$V_\tau \;=\; \mathcal{L}V + \max_{\substack{\varphi\in\{0,1\},\psi\in\{0,1\}\\ \varphi\psi=0}} \left[\varphi G_r \mathcal{F}V + \psi\left(\frac{(\mathcal{F}V-\kappa)}{\varepsilon} + \kappa G_r\right)\right] . \tag{6}$$

Basic Idea:

- Discretize equation (6)
- Let $\varepsilon \to 0$ as the mesh and timesteps tend to zero
- We will show that a discrete version of (6) is consistent in the viscosity sense with the singular control problem (1)

Previous Work

- Milevsky, Salisbury, (2006, Insurance: Mathematics and Economics), pose GMWB pricing problem as a singular control.
- Dai, Kwok, Zong, (Mathematical Finance, 2008), propose the penalty method to solve the singular control formulation of the GMWB
- Chen, Forsyth (Numerische Mathematik, 2008), solve impulse control formulation of the GMWB

Localization

The original GMWB problem is posed on the domain

$$(W, A, \tau) \in [0, \infty] \times [0, \omega_0] \times [0, T] \ . \tag{7}$$

where ω_0 is the initial contract premium.
For computational purposes, we localize this domain to

$$\Omega \ = \ [0, W_{\text{max}}] \times [0, \omega_0] \times [0, T] \ . \tag{8}$$

Apply boundary conditions (see Dai et al (2008)).

Assumption

The localized GMWB singular control problem satisfies a strong comparison result, hence a unique continuous viscosity solution exists.

Remark

In (Seydel (2008)), it is shown that an impulse control formulation of the GMWB pricing problem satisfies a strong comparison principle. However, there does not seem to be a proof of this result for the singular control formulation of this problem.

Discretization

Define a set of nodes

$$\{W_0, \ldots, W_i, \ldots, W_{i_{\max}}\} \quad ; \quad \{A_0, \ldots, A_j, \ldots, A_{j_{\max}}\} \tag{9}$$

Let $V_{i,j}^n$ be the approximation of $V(W_i, A_j, \tau^n)$.

Define vector of nodal values along lines of constant A_j

$$\vec{V}_j^n = \left[V_{0,j}^n, V_{1,j}^n, \ldots, V_{i_{\max},j} \right]' \tag{10}$$

Fully implicit timestepping

- Use forward, backward central differencing
- Use central differencing as much as possible yet still retain a *positive coefficient* scheme (Wang and Forsyth (2008)).

Discretization

We can write this in matrix form

$$\left[I + \Delta\tau M_j^{n+1} \right] \vec{V}_j^{n+1} = \vec{V}_j^{n} + \Delta\tau P_j^{n+1} \vec{V}_{j-1}^{n+1} + \Delta\tau \vec{D}_j^{n+1}. \qquad (11)$$

Note that $M_j^{n+1}, P_j^{n+1}, \vec{D}_j^{n+1}$ are functions of the optimal controls $\{\varphi_{i,j}^{n+1}, \psi_{i,j}^{n+1}\}$.

Let $\mathcal{F}_h V_{i,j}^n$ be the discrete form of $\mathcal{F}V = 1 - V_W - V_A$.

The optimal controls at each node are given by

$$\{\varphi_{i,j}^{n+1}, \psi_{i,j}^{n+1}\} \in \underset{\substack{\varphi\in\{0,1\},\psi\in\{0,1\} \\ \varphi\psi=0}}{\arg\max} \left[\varphi G_r \mathcal{F}_h V_{i,j}^{n+1} + \psi \left(\frac{(\mathcal{F}_h V_{i,j}^{n+1} - \kappa)}{\varepsilon} + \kappa G_r \right) \right]$$

$\rightarrow$ This makes equation (11) highly nonlinear.

Positive Coefficient Scheme

Recall our discretization in matrix form

$$\left[I + \Delta\tau M_j^{n+1}\right] \vec{V}_j^{n+1} = \vec{V}_j^n + \Delta\tau P_j^{n+1} \vec{V}_{j-1}^{n+1} + \Delta\tau \vec{D}_j^{n+1}.$$

If a positive coefficient method is used, then $\left[I + \Delta\tau M_j^{n+1}\right]$ is an $\mathbb{M}$ matrix

- Recall that this means that $\left[I + \Delta\tau M_j^{n+1}\right]^{-1} \geq 0$.

We use a type of Policy iteration to solve the discretized nonlinear equations

Theorem (Convergence of the Policy Iteration)

If the positive coefficient condition is satisfied, then the policy iteration converges to the unique solution of (11) for any initial estimate $(\vec{V}_j^{n+1})^0$.

Proof.

Same steps as in (Forsyth and Labahn (2008)) □

Stability, Monotonicity

Lemma (Stability)

If the discretized scheme satisfies the positive coefficient condition, the method is l_∞ stable.

Proof.

Use maximum analysis and induction. □

Lemma (Monotonicity)

If the discretized scheme satisfies the positive coefficient condition, the method is unconditionally monotone.

Proof.

Similar steps as in (Forsyth and Labahn (2008)). □

Remark (Monotonicity)

Monotonicity can be viewed as a discrete arbitrage inequality, i.e. inequality of payoffs implies inequalities of value at all earlier times.

Consistency

Assume

$$h = \frac{\Delta W_{\max}}{C_1} = \frac{\Delta A_{\max}}{C_2} = \frac{\Delta \tau}{C_3} = \frac{\varepsilon}{C_4}, \tag{12}$$

Lemma (Consistency)

In the limit as $h \to 0$, the numerical scheme is consistent, in the viscosity sense (Barles, Souganidis (1991)), with the singular control problem

$$\min\left[V_\tau - \mathcal{L}V - G_r \max(\mathcal{F}V, 0), \kappa - \mathcal{F}V\right] = 0$$

Proof.

Mimic steps used in the informal derivation of the singular control problem, only this time, with finite difference approximations of the operators and smooth test functions. □

Convergence

Theorem

Provided that (a) the GMWB singular control problem satisfies a strong comparison result (b) fully implicit timestepping is used with a positive coefficient discretization, then the penalty scheme converges to the unique, continuous viscosity solution of the GMWB singular control problem.

Proof.

The scheme is monotone, consistent and stable, hence this follows from the results in (Barles, Souganidis (1991)). □

Remark

Note that convergence is guaranteed for any penalty term ε *of the form* $\varepsilon = Ch$*, where* h *is the discretization parameter.*

Examples

Recall that the investor pays no extra up-front fee for the guarantee (only the initial premium w_0).

The insurance company deducts an annual fee η from the balance in the sub-account W.

Problem: let $V(\eta, W, A, \tau)$ be the value of the GMWB contract, for given yearly guarantee fee η.

Assume that the investor pays an initial premium w_0 at $t = 0$ ($\tau = T$).

Find the no-arbitrage fee η such that $V(\eta, w_0, w_0, T) = w_0$ (we do this by a Newton iteration).

Penalty Parameter

We obtain convergence if

$$\varepsilon = C\Delta\tau$$

for any $C > 0$.

Recall that the maximum withdrawal rate is $\lambda = 1/\varepsilon$

- If $\lambda = \omega_0/(\Delta\tau) \rightarrow$ entire guarantee amount can be withdrawn in one timestep
  $\hookrightarrow$ Effectively an infinite rate
- This suggests that a reasonable value for $\varepsilon = 1/\lambda$ would be

$$\begin{aligned} \varepsilon &= \Delta\tau C^*/\omega_0 \\ & C^* = \text{ dimensionless constant } < 1 \end{aligned}$$

- But we also want to make $\varepsilon\left(V_\tau - \mathcal{L}V - \kappa G_r\right)$ small on coarse grids.
- We choose $C^* = 10^{-2}$. More later.

Example: Data

Parameter	Value
Expiry time T	10.0 years
Interest rate r	.05
Maximum withdrawal rate G_r	10/year
Withdrawal penalty κ	.10
Volatility σ	.30
Initial Lump-sum premium w_0	100
Initial guarantee account balance	100
Initial sub-account value	100
Penalty parameter ε	$10^{-2}\Delta\tau/\omega_0$

Grid/Timesteps

Level	W Nodes	A Nodes	Time steps
1	117	111	120
2	233	221	240
3	465	441	480
4	929	881	960
5	1857	1761	1920

Table: Grid and timestep data for convergence experiments

Results: No-arbitrage Fee Charged

Refine	Central Differencing First			Backward Differencing Only		
Level	Value	Itns/step	Ratio	Value	Itns/step	Ratio
Fully Implicit Method						
$\sigma = 0.2, \eta = 0.013886$						
1	101.3329	3.51	N/A	101.7167	3.48	N/A
2	100.4556	3.62	N/A	100.7628	3.52	N/A
3	100.1283	3.70	2.68	100.3210	3.63	2.15
4	100.0271	3.77	3.23	100.1284	3.71	2.29
5	100.0000	3.89	3.71	100.0439	3.87	2.28
$\sigma = 0.3, \eta = 0.031286$						
1	100.5987	4.19	N/A	100.9427	4.09	N/A
2	100.1495	4.31	N/A	100.3602	4.26	N/A
3	100.0358	4.33	3.95	100.1280	4.31	2.51
4	100.0081	4.39	4.11	100.0472	4.38	2.88
5	100.0000	4.38	3.40	100.0183	4.37	2.80

- Value at $S = A = 100, t = 0$.

Why do we get almost second order convergence?

Several lucky breaks

Fully implicit timestepping should only be first order

- $V_{\tau\tau} \simeq 0$ as $\tau \rightarrow T$.
- Fully implicit timestepping is exact

In infinite rate withdrawal regions:

- Backward differencing used
- But solution in these regions $V \simeq$ linear in $(W, A) \rightarrow$ backward differencing is exact

Penalty Parameter

Recall the penalty parameter ε

$$V_\tau = \mathcal{L}V + \max_{\substack{\varphi \in \{0,1\}, \psi \in \{0,1\} \\ \varphi\psi = 0}} \left[\varphi G_r \mathcal{F} V + \psi \left(\frac{(\mathcal{F}V - \kappa)}{\varepsilon} + \kappa G_r \right) \right] .$$

The penalty method will converge for any penalty parameter ε of the form

$$\varepsilon = C \Delta\tau$$

Based on financial reasoning, a good choice for ε is

$$\varepsilon = \frac{C^* \Delta\tau}{\omega_0} \quad ; \quad C^* \leq 1$$

where C^* is dimensionless.

We would like to choose C^* small, so that convergence is rapid, but if C^* is too small $\rightarrow$ machine precision problems

How do we choose C^*?

Penalty Parameter: tests

	$\sigma = 0.2$		$\sigma = 0.3$	
ε	Value	Itns/step	Value	Itns/step
$\Delta\tau/\omega_0$	107.7315	3.3	115.8828	3.2
$10^{-1}\Delta\tau/\omega_0$	107.7336	3.2	115.8856	3.3
$10^{-2}\Delta\tau/\omega_0$	107.7338	3.2	115.8859	3.3
$10^{-3}\Delta\tau/\omega_0$	107.7339	3.2	115.8860	3.3
$10^{-4}\Delta\tau/\omega_0$	107.7339	3.3	115.8860	3.3
$10^{-5}\Delta\tau/\omega_0$	107.7339	3.2	115.8860	3.3
$10^{-6}\Delta\tau/\omega_0$	107.7339	3.2	115.8860	3.3
$10^{-7}\Delta\tau/\omega_0$	107.7338	3.3	115.8860	3.3
$10^{-8}\Delta\tau/\omega_0$	107.7336	3.4	115.8854	3.5
$10^{-9}\Delta\tau/\omega_0$	107.7244	4.9	115.8615	5.4

Table: The effect of the penalty parameter at refinement level 5. $W = A = 100$ and $t = 0$. No insurance fee (i.e. $\eta = 0$) is imposed.

- Any $C^* \in [10^{-1}, 10^{-7}]$ works fine.

Optimal Withdrawal Strategy

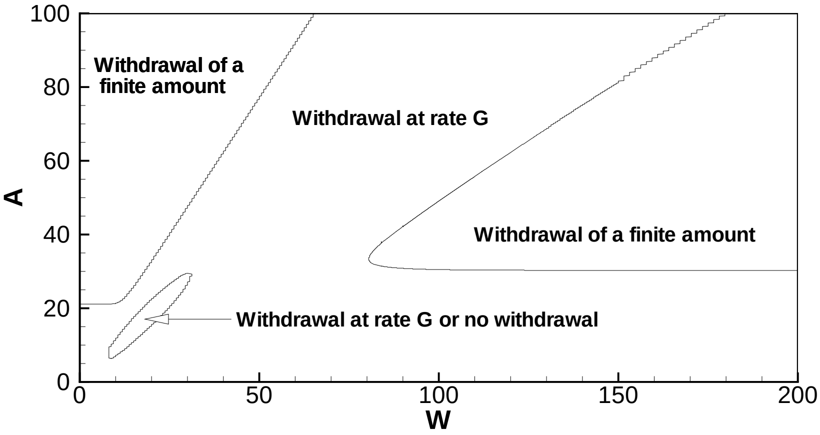

Figure: The contour plot of optimal withdrawal strategy of the GMWB guarantee at $t = 0$ in the (W, A)-plane. $\sigma = 0.3$. Fair fee $\eta = .031286$ is imposed.

No-arbitrage Fee

- $\sigma = .15 \rightarrow \alpha = .007$ (70 bps)
- $\sigma = .20 \rightarrow \alpha = .014$ (140 bps)
- $\sigma = .30 \rightarrow \alpha = .031$ (310 bps)
- Current volatility of $S\&P =?$
- Typical fees charged: $\alpha = .005$ (50 bps) too low for current market conditions.
- Insurance companies seem to be charging fees based on marketing considerations, not hedging costs.
- Fee should be even higher if other (typical) contract options considered (sales fees), and other processes (jumps, see Chen, Forsyth, Ins. Math. Econ. (2008)).

Comparison with Impulse Control Formulation

Chen and Forsyth (2008) solve an impulse control formulation

$$\min\Big\{ V_\tau - \mathcal{L}V - \max_{\gamma \in [0, G_r]} \big(\gamma - \gamma V_W - \gamma V_A\big),$$

$$V - \sup_{\hat{\gamma} \in (0, A]} \big[V(\max(W - \hat{\gamma}, 0), A - \hat{\gamma}, \tau) + (1-\kappa)\hat{\gamma} - c \big] \Big\} = 0$$

where c is a small (infinitesimal) fixed cost.

If h is the mesh size parameter, then

- Complexity of the Impulse Control solution is $O(h^{-4})$
- This is due to the linear search required for solving the local optimization problems at each node
- Compare with complexity of $O(h^{-3})$ for the penalized singular control formulation
- But the Impulse Control formulation is more general, i.e. some contract features cannot be modelled with the Singular Control formulation

Summary

- The Penalty method can be proven to converge to the viscosity solution of the Singular Control formulation of the GMWB problem (assuming a strong comparison property holds).
- The Penalty method is very simple to implement, results not sensitive to choice of dimensionless constant in the penalty term.
- Easy to apply same idea to other singular control problems.
- Lower complexity than the Impulse Control Formulation
  - $\hookrightarrow$ But Impulse Control is more general
- If penalty parameter too small $\rightarrow$ roundoff problems
  - $\hookrightarrow$ But not a problem of practical concern, i.e. eight digit accuracy obtained without difficulty
- Insurance companies seem to be charging fees which are too low to cover hedging costs. Mark-to-market writedowns already occurring in Canada and US.