# A Primer on the Mathematics of Guaranteed Minimum Maturity Benefit (GMMB) Insurance Contracts

Haoen Cui
hcui10@illinois.edu

May 11, 2018

**Abstract**

This primer is the result of Haoen Cui's work in MATH 492 (Spring 2018) under supervision of Prof. Runhuan Feng. Many of the contents are already included in [1] as examples. The goal of this primer is to introduce concepts surrounding equity-linked life insurance products through a GMMB example to readers with introductory proficiency in undergraduate-level actuarial topics. For a more comprehensive reference, one should consult [1].

**Key Words.** GMMB; Black-Scholes PDE; Risk-Neutral Pricing; Hedging; Linear Interpolation; Least-Square Monte Carlo; Taylor's Expansion; Linear Regression.

# Contents



# List of Figures



# List of Tables

# 1 Introduction

As its name suggests, **equity-linked life insurance** is a special class of life insurance policies whose benefit (to the insured) is defined with respect to the market value of the contract-determined reference portfolio, hence equity-linked. One can further classify such policies into *pure equity-linked contracts*, when the policyholder bears all financial risks, and *guaranteed equity-linked contracts*, where both the insured and the insurer share market risks. Feng's forthcoming book [1] covers many such products in the second category which are often referred to as **GMXB**, such as

- Guaranteed Minimum Maturity Benefit (GMMB),
- Guaranteed Minimum Death Benefit (GMDB),
- Guaranteed Minimum Accumulation Benefit (GMAB), and
- Guaranteed Minimum Withdrawal Benefit (GMWB).

This paper will focus on the first case, **GMMB**. We will analyze the financial aspects of this type of contract under the Black-Scholes framework. Introductory knowledge on models of life contingencies, statistics, and stochastic calculus is assumed.

The rest of the paper is organized as follows: Section 2 discusses GMMB contract mechanisms and derives a pricing formula; Section 3 expands on hedging which includes continuously hedging with exogenous cash flows and greeks letters hedging; Section 4 focuses on two interpolation techniques, linear interpolation and least-square Monte Carlo method; Section 5 concludes the paper and provides avenues for future work.

# 2 Pricing and Valuation

## 2.1 GMMB Contract Mechanics

Without taping into business details, interactions between the policyholder and the insurer under a GMMB contract can be simplified to the model below.

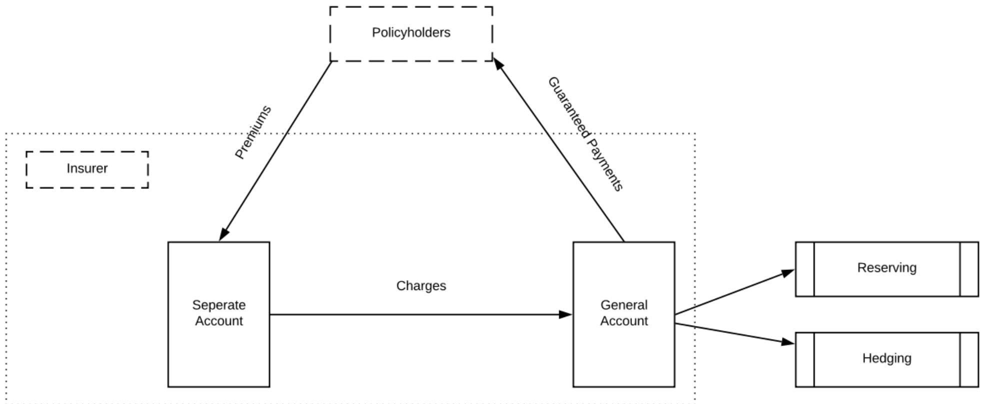


Figure 1: Simplified Diagram of Cash Flows for GMMB Contracts

We assume that policyholders pay a one-time gross premium to the their separate account at the time of contract issuance, which will accumulate value due to investment returns. During the time when the policy is still in effective, the insurer charges periodic asset-value-based fees from policyholders' sub-accounts. Upon policyholder's death, the insurer will pay the death benefit which is the larger of (1) the market value of the policyholders sub-account or (2) the pre-determined guarantee amount. For simplicity of our discussion, we assume charges are transfered to the insurer's general account at the same time of valuation periods and the death benefit is payable precisely at the time of death. In particular, we sill study how an insurer will *value* a contract, and how it will *hedge* using the funds available through its general account.

## 2.2 Notations

We will use the same notations as defined in [1].

- $S_t$, the market value of the underlying equity (portfolio) index at $t \geq 0$.
- $F_t$, the market value of the policyholders sub-accounts at $t \geq 0$. $F_0$ is considered to be the initial premium invested at the start of the contract.
- $G$, the guaranteed minimum death benefit amount. Note that $G$ is a constant in our discussion for GMMB contracts.

- $n$, the number of valuations per year.

- $m$, the nominal annualized rate at which asset-value-based fees are deducted from sub-accounts. Note that [1] denotes the annualized rate of charges allocated to the GMMB by $m_e$ whenever there is a need to distinguish from other contracts.

- $r$, the continuously compounding annual risk-free rate. Note that $r$ is a constant in our discussion.

- $T$, the maturity date of the insurance policy.

- $T_x$, the future lifetime of the policyholder of age $x$ at inception.

- ${}_tp_x$, the probability that a living individual at age $x$ will survive for another $t$ years

## 2.3 Net Present Value of Future Liabilities

As a common practice in life contingency modeling, we consider the *net present value random variable* of the GMMB future liabilities at the start of the contract with $n$ valuation periods per year,

$$L^{(n)}(T_x) := e^{-rT}(G - F_T)_+ \mathbb{1}(T_x > T) - M^{(n)}_{T \wedge T_x} \tag{1}$$

where $(x)_+ := \max\{x, 0\}$ and $x \wedge y := \min\{x, y\}$. $M^{(n)}_{T \wedge T_x}$ in the equation is called *margin offset* with $n$ valuation periods per year, which is given by

$$M^{(n)}_t := \sum_{j=1}^{\lceil nt \rceil} e^{-r\frac{j-1}{n}} \cdot \frac{m}{n} \cdot F_{\frac{j-1}{n}} \tag{2}$$

The equity-linked mechanism for variable annuity dictates that at the end of each trading day, the account value fluctuates in proportion to the value of equity fund in which it invests and deducted by account-value-based fees. With $n$ equidistant valuation periods per year, policyholder's sub-account value at the end of each sub-period is given by

$$F_{\frac{k}{n}} = F_0 \frac{S_{\frac{k}{n}}}{S_0} \left(1 - \frac{m}{n}\right)^k \qquad k = 1, 2, \cdots, nT \tag{3}$$

Equipped with the widely known limit, $\lim_{n\to+\infty}(1-\frac{m}{n})^n = e^{-m}$, we can shrink the valuation period to zero and derive the continuous-time expressions

$$L^{(\infty)}(T_x) := e^{-rT}(G - F_T)_+\mathbb{1}(T_x > T) - \int_0^{T\wedge T_x} e^{-rs}mF_s ds \tag{4}$$

$$M_t^{(\infty)} = \int_0^t e^{-rs}mF_s ds \tag{5}$$

$$F_t = \frac{F_0}{S_0}S_t e^{-mt} \tag{6}$$

From now on, we will drop the $(\infty)$ superscript and only keep the $t$ subscript to denote the continuous-time quantities without ambiguities.

## 2.4 The Black-Scholes Framework

To the policyholder, the payoff from a GMMB contract is very similar to that of a put option. It makes sense for us to recall the Black-Scholes framework from option pricing context. We assume that the underlying asset $S_t$ follows a geometric Brownian motion

$$\frac{dS_t}{S_t} = (\mu - \delta)dt + \sigma dW_t \tag{7}$$

where the instantaneous return $\mu$, instantaneous dividend rate $\delta$, and the instantaneous volatility $\sigma$ are assumed to be constants. $dW_t$ is often referred as a standard pure Brownian motion or Wiener process. It can be shown that the asset price follows a log-normal distribution at any time $t$

$$S_t \sim S_0 \cdot \text{LogNormal}\big((\mu - \delta - \frac{1}{2}\sigma^2)t, \sigma^2 t\big) \tag{8}$$

Hence its descriptive statistics are given by

$$\mathbb{E}[S_t] = S_0 e^{(\mu-\delta)t} \qquad \text{Var}[S_t] = {S_0}^2 e^{2(\mu-\delta)t}(e^{\sigma^2 t} - 1)$$

The value $V$ of any option contract on $S_t$ is given by a functional $f(S, t; \sigma; G, T; r)$ where semi-colons separate different types of variables and parameters,

- $S$ (asset price) and $t$ (time) are variables

- $\sigma$ (volatility) is a parameter associated with underlying asset
- $G$ (strike price) and $T$ (maturity date) are parameters associated with the contract configurations
- $r$ (risk-free interest rate) is a parameter associated with the currency

We often drop all parameters and simply write $f(S,t)$. Notice that $f$ does not depend on $\mu$, the instantaneous rate of return of the underlying asset, which is a result of the Black-Scholes model. This model argues that one can construct a hedged portfolio of the form

$$H(S,t) = f(S,t) - \Delta(S,t) \cdot S \tag{9}$$

at any time $t$ and any asset price $S$ such that the portfolio is risk-free. To see this, we compute $dV$ based on *Ito's Lemma*,

$$df = \left(\frac{\partial f}{\partial t} + \mu S \frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2}\right) dt + \sigma S \frac{\partial f}{\partial S} dW$$

and plug into $dH = df - \Delta dS - \Delta \delta S dt$,

$$dH = \left(\frac{\partial f}{\partial t} + \mu s\Big(\frac{\partial f}{\partial S} - \Delta\Big) + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} - \Delta \delta S\right) dt + \sigma S\Big(\frac{\partial f}{\partial S} - \Delta\Big) dW$$

We can easily see that if

$$\Delta(S,t) = \frac{\partial f(S,t)}{\partial S} \tag{10}$$

then $H$ is risk-free, hence it must be compounded at a risk-free return, i.e. $dH = rHdt$, which leads to

$$\left(\frac{\partial}{\partial t} + (r-\delta) S \frac{\partial}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2} - r\right) f = 0 \tag{11}$$

This PDE is known as the *Black-Scholes Equation*, and it is a linear parabolic PDE. We can identify the terms in the parentheses as the *Black-Scholes Operator*. Financially, it means that if we can continuously re-balance (a.k.a. dynamically hedge) our hedging portfolio $H$, then the value function $f$ is given by the solution of the above PDE with appropriate boundary conditions according to the contract configurations.

## 2.5 Risk-Neutral Pricing

One way to find the solution is using *Feymann-Kac Formula* and obtaining the solution in the form of conditional expectation. In fact, the theorem covers a broader class of PDEs than just the Black-Scholes PDE. One of its consequence is that the no-arbitrage price at time $t$ of any path-independent financial derivative (e.g., $B(t,F)$) with maturity at time $T$ is given by a conditional expectation taken with respect to a *risk neutral measure* $\mathbb{Q}$ where any financial derivative would have the risk-free rate as their average rate of return.

$$\text{Under } \mathbb{Q} \qquad \frac{dS_t}{S_t} = (r-\delta)dt + \sigma dW_t \tag{12}$$

An analogue can also be derived to price path-dependent derivatives (e.g., $P(t,F)$). One shall consult [1] for details. We will only state the results on the valuation of net present value of a GMMB contract's future liabilities

$$L_t = \underbrace{e^{-r(T-t)}(G-F_T)_+\mathbb{1}(T_x > T)}_{\text{Benefit Portion}} - \underbrace{\int_t^{T\wedge T_x} e^{-rs} m F_s ds}_{\text{Payment Portion}}$$

where the value of benefit portion payable to the insured contingent on survival beyond maturity is

$$B(t,F) = {}_Tp_x\Big[Ge^{-r(T-t)}\Phi\Big(-d_2\big(T-t,\frac{F}{G})\Big) - Fe^{-m(T-t)}\Phi\Big(-d_1\big(T-t,\frac{F}{G})\Big)\Big] \tag{13}$$

$$\begin{aligned} d_1\big(T-t,\frac{F}{G}) &= \frac{\ln(\frac{F}{G}) + (r-m+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ d_2\big(T-t,\frac{F}{G}) &= d_1\big(T-t,\frac{F}{G}) - \sigma\sqrt{T-t} \end{aligned}$$

and the payment portion receivable by the insurer until the earlier date of maturity or the insured's death is

$$P(t,F) = mF\int_t^T e^{-m(s-t)}{}_sp_x ds \tag{14}$$

Not surprisingly, expression of $B(t,F)$ is a survival probability times a put option on $F$, the policy holder's sub-account value, with a strike price of guaranteed minimum maturity benefit $G$ and fee rate $m$ as "dividend." We will denote the net present value of a GMMB contract to its policy holder by $V$ and identify its two components,

$$V(t, F) = B(t, F) - P(t, F) \tag{15}$$

The *fair rate* $m^*$ is determined such that the value $V$ is zero at issuance. Hence, we equate $B(t = 0, F_0) = P(t = 0, F_0)$ to solve for $m^*$. Unless otherwise mentioned, we will simply use $m$ to denote the fair rate $m^*$.

# 3 Risk Management via Hedging

## 3.1 Continuously Hedging with Exogenous Cash Flows

As alluded to in the Black-Scholes framework, the valuation is based on the assumption of continuously hedging via constructing the portfolio

$$\underbrace{H(S_t, t)}_{H_t} = \underbrace{V(S_t, t)}_{V_t} - \underbrace{\Delta(S_t, t)}_{\Delta_t} \cdot S_t \quad \text{where} \quad \Delta_t = \left.\frac{\partial V(S, t)}{\partial s}\right|_{S=S_t}$$

(Note that we consider $S_t$ instead of $F_t$ because $S_t$ is a tradable underlying asset but $F_t$ is not.) However, the Black-Scholes PDE is not satisfied by the GMMB valuation formula $V$ because equity-linked annuity products are often structured with intermediate cash flows over their policy terms. In fact,

$$\left(\frac{\partial}{\partial t} + rS\frac{\partial}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2} - r\right) V(t, F(S)) = -{}_t p_x m F \tag{16}$$

Therefore, we will extend the idea of self-financing hedging portfolio to incorporate exogenous cash flows. Consider a replicating portfolio $H_t^*$ in the form of

$$H_t^* := h_t S_t + B_t^* \tag{17}$$

where $h_t$ denotes the number of shares invested in the underlying asset and $B_t^*$ is the value of the money market account. Its dynamics is given by

$$dH_t^* = h_t dS_t + rB_t^* dt + C_t dt \tag{18}$$

where $C_t$ is the annualized rate of net cash flow (incomes less outgoes). We will let $H_t^*$ to match $V_t$ and $dH_t^*$ to match $dV_t$, which leads to

$$h_t = \frac{\partial V(t, S_t)}{\partial S} = \Delta_t \tag{19}$$

and

$$C_t = \left(\frac{\partial}{\partial t} + rS\frac{\partial}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2} - r\right)V(t, F(S_t)) = -{}_tp_x m F_t \tag{20}$$

Moreover, there is no upfront fee for GMMB contracts which makes $H_t^*$ a "self-financing" portfolio with no initial investment.

## 3.2 Discrete Time Approximation

In real life scenario, we cannot trade continuously, hence need discrete-time approximation to the above results. Assume that we have at discrete-time mesh points $\{t_n\}_{n=0}^N \subset [t, T]$ with equidistant interval length $\Delta t$. We denote the re-balanced portfolio after the $n$-th hedge by

$$H_n^* := \Delta_n^* S_n + B_n^* \tag{21}$$

then before the next hedge

$$\begin{aligned} H_{n+1} \quad &= \quad \Delta_n^* S_{n+1} + e^{r\Delta t} B_n^* + \int_{t_n}^{t_{n+1}} C_t dt \\ &\overset{\text{match}}{=} \quad H_{n+1}^* \quad := \Delta_{n+1}^* S_n + B_{n+1}^* \end{aligned}$$

Thus, money market account balance after hedging can be determined recursively

$$\begin{aligned} B_{n+1}^* &= e^{r\Delta t} B_n^* - (\Delta_{n+1}^* - \Delta_n^*) S_{n+1} + \int_{t_n}^{t_{n+1}} C_t dt && (22) \\ &= e^{r\Delta t} B_n^* - (\Delta_{n+1}^* - \Delta_n^*) S_{n+1} - \int_{t_n}^{t_{n+1}} {}_tp_x m F_t dt && (23) \end{aligned}$$

### 3.2.1 Numerical Example

Intuitively, the smaller the time interval $\Delta t$ is, the more accurate our replication $H_n^*$ will match against the true analytic value $V_t$. In addition, accuracy also depends other factors such as method used for numerical quadrature $\int_{t_n}^{t_{n+1}} {}_tp_x m F_t dt$ and models for life contingency ${}_tp_x$. In this section, we provide a numerical example using the `integrate` function from base `R` package and a life table with UDD assumption for fractal age.

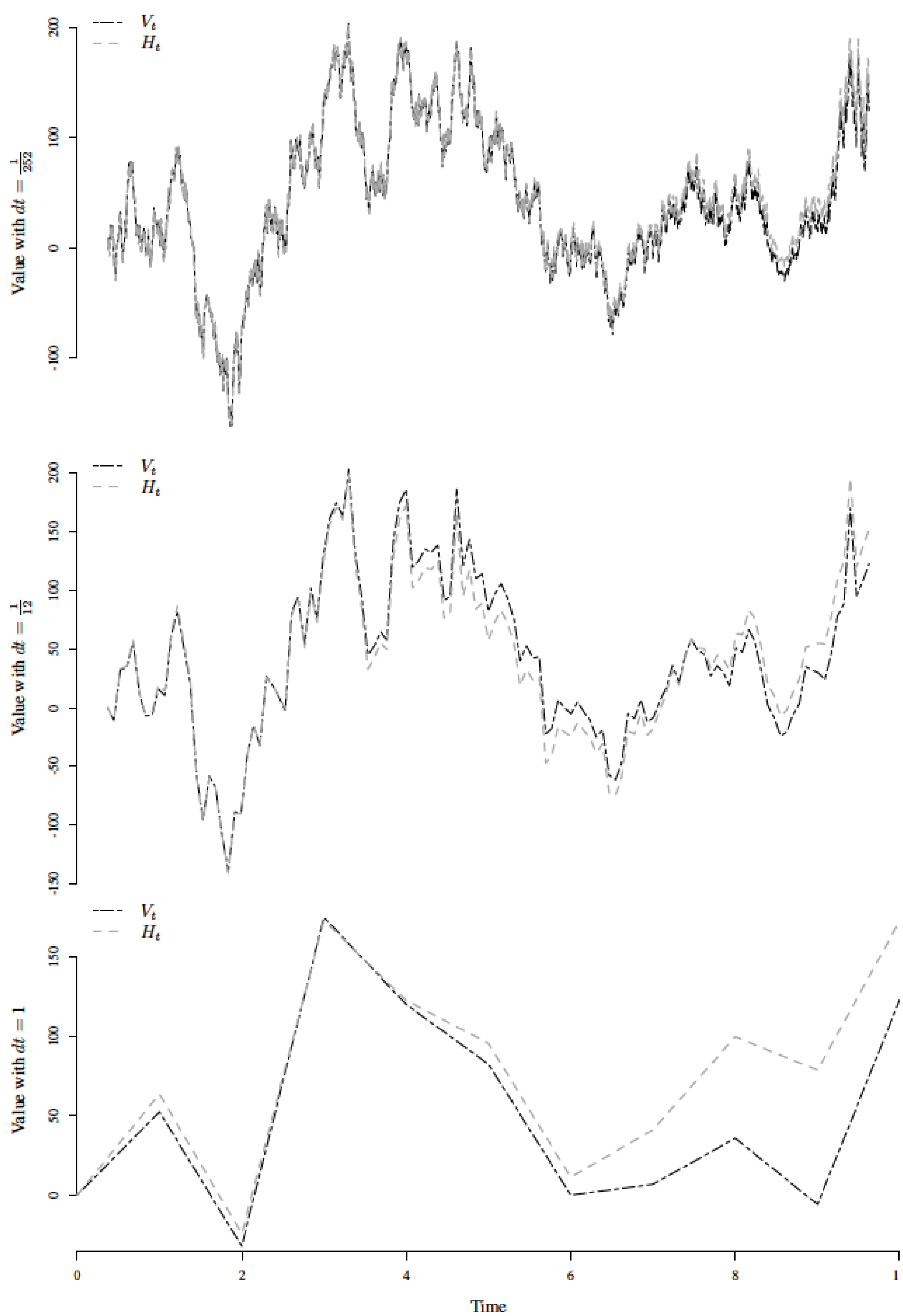


Figure 2: Pathwise Comparison of Replicating Portfolio Value with Different Rebalancing Frequencies

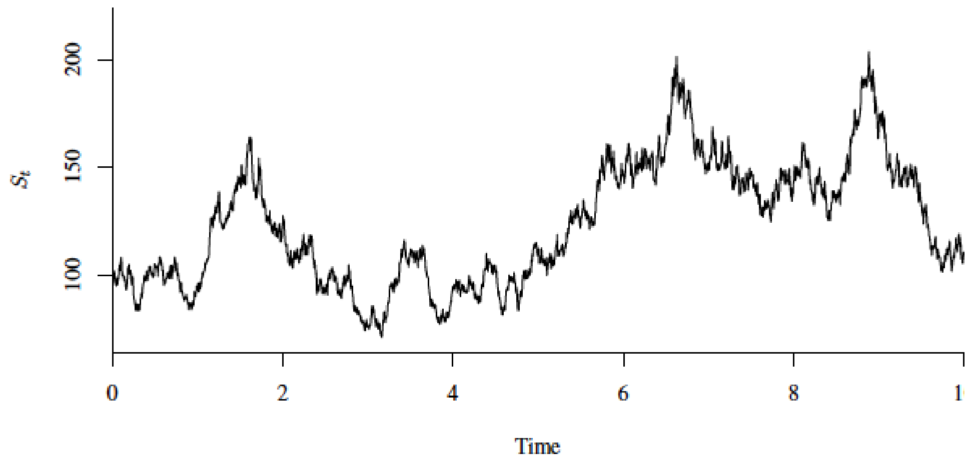


Figure 3: Sample path of Underlying Asset Price

Above, we provided plots for a simulated price path of the underlying asset $S_t$ and the replicating results $V_t$ v.s. $H_t^*$ with different hedging frequencies $\Delta t \in \{1, \frac{1}{12}, \frac{1}{252}\}$ (yearly, monthly, and trading-daily). One can easily see that

- Yearly re-balancing has quite large tracking errors;
- Monthly re-balancing has reasonably good performance in this sample;
- One can hardly distinguish trading-daily hedging from the corresponding true values.

In practice, frequent re-balance will incur huge trading costs. One may need to trade off between the benefit of closer tracking and additional costs due to trading.

## 3.3 Greeks Letter Hedging

"All models are wrong", generally attributed to statistician George Box, is a common aphorism in many applied mathematical modeling. Many inputs that we assumed to be constant parameters, such as $r$ and $\sigma$, can very much change in reality. In fact, one common practice is to approximate the portfolio's value via Taylor's expansion over inputs $\mathbf{x} = (x^{(i)})_i$ that are measurable or tradable in the market.

$$f(\mathbf{x}+\Delta\mathbf{x})-f(\mathbf{x})\approx\sum_i\frac{\partial}{\partial x^{(i)}}f(\mathbf{x})\Delta x^{(i)}$$

Some partial derivatives so widely-used that they deserve to have their own names. These partial derivatives are collectively known as *greek letters*. For example, $\Delta(t,S_t)$ is arguably one of the most famous ones among them. Hedging against a collection of greek letters is known as *greek letters hedging*.

| **Greek** | **Definition** | **Expression** |
|---|---|---|
| $\Delta$ | $\frac{\partial}{\partial S}V$ | $-{}_Tp_x\frac{F}{S}e^{-m(T-t)}\Phi(-d_1(T-t,\frac{F}{G}))-\frac{P}{S}$ |
| $\Gamma$ | $\frac{\partial^2}{\partial^2 S}V$ | ${}_Tp_x\frac{F^2}{S^2}e^{-m(T-t)}\frac{\phi\big(d_1(T-t,\frac{F}{G})\big)}{F\sigma\sqrt{T-t}}$ <br> $={}_Tp_x\frac{F^2}{S^2}Ge^{-r(T-t)}\frac{\phi\big(d_2(T-t,\frac{F}{G})\big)}{F^2\sigma\sqrt{T-t}}$ |
| $\mathcal{V}$ | $\frac{\partial}{\partial\sigma}V$ | ${}_Tp_xFe^{-m(T-t)}\phi\big(d_1(T-t,\frac{F}{G})\big)\sqrt{T-t}$ <br> $={}_Tp_xGe^{-r(T-t)}\phi\big(d_2(T-t,\frac{F}{G})\big)\sqrt{T-t}$ |
| $\Theta$ | $\frac{\partial}{\partial t}V$ | ${}_Tp_x\cdot\Bigg[rGe^{-r(T-t)}\Phi\Big(-d_2(T-t,\frac{F}{G})\Big)$ <br> $-mFe^{-m(T-t)}\Phi\Big(-d_1(T-t,\frac{F}{G})\Big)$ <br> $-e^{-m(T-t)}\frac{\sigma F\phi\big(d_1(T-t,\frac{F}{G})\big)}{2\sqrt{T-t}}\Bigg]+{}_tp_xmF$ |

Table 1: GMMB Greeks

Formally, if we have a collection of tradable contracts with values prescribed by $(g_j)_j$, we would like to find corresponding holdings $(h_j)_j$ such that after combining with our current portfolio $f$, the resulting portfolio will not have any exposure to a set of $I$ pre-determined greeks, i.e.

$$\frac{\partial}{\partial x^{(i)}}(f+\sum_j g)=0\qquad\forall i\in[I]\tag{24}$$

We can also express the criteria in matrix form. Consider greeks matrix $A=(A_{ij})_{ij}$ where components $A_{ij}$ represents the value of $j$-th contract's

$i$-th greeks. Let vector $\mathbf{b} = (b_i)_i$ denotes that greeks values of our current portfolio $f$. Then, we aim to find holdings $\mathbf{x} = (x_i)_i$ such that

$$A\mathbf{x} = \mathbf{b} \tag{25}$$

One can apply any acceptable rounding conventions to the solution of holdings if contracts can not be traded in fractions.

### 3.3.1 Sherman-Morrison Update

A computational trick that may be helpful is the *Sherman-Morrison formula*. It provides a numerically cheap way to compute the inverse of a matrix $A$ corrected by a perturbation $uv^\top$ if the inverse of $A$ is already known. Precisely, suppose $A \in \mathbb{R}^{n\times n}$ is an invertible matrix and $u, v \in \mathbb{R}^n$ are column vectors, then $A + uv^\top$ is invertible if and only if $1 + v^\top A^{-1}u \neq 0$. In addition, if $A + uv^\top$ is invertible, then its inverse is given by

$$(A + uv^\top)^{-1} = A^{-1} - \frac{A^{-1}uv^\top A^{-1}}{1 + v^\top A^{-1}u} \tag{26}$$

In the general case, if $u$ and $v$ are arbitrary vectors, then the computation tales $3n^2$ scalar multiplications, where there is benefit. However, if either $u$ or $v$ (but not both) is a unit vector, then only $2n^2$ multiplications are needed. Furthermore, if both are unit vectors, then only $n^2$ calculations are necessary. This result can be helpful when we want to test a trading strategy over time is only a small number of greeks change frequently. For example, we know that the underlying asset $S_t$ has constant $\Delta$ and no exposure to any other greeks.

# 4 Interpolation Techniques in Nested Stochastic Modeling

In *nested stochastic modeling*, there are typically two layers of simulation, the outer loop and the inner loop. As their name suggests, the outer loop scenarios of a given risk factor and the inner loop simulates paths for another risk factor for each scenario in the outer loop. By definition, the complexity is exponential which limits its usage in practice. Therefore, people consider *preprocessed inner loop* (a.k.a. *factor-based approach*) to save computation time. The essence is to run inner loop only on a small set of outer-loop scenarios (hence "preprocessing") and interpolate the preprocessed calculation results to the entire domain of interest.

In this section, we review the theory of two commonly used methods. For simplicity, we will only consider a two-dimensional case. Let $(x, y)$ be the pair of underlying risk factors, and suppose we have preprocessed values of $\hat{L}(x_i, y_j)$ on mesh points $\{(x_i, y_j) : i = 1, 2, \cdots, m; j = 1, 2, \cdots, n\}$. Now, we will apply interpolation to approximate $\hat{L}(x, y)$ for any $(x, y) \in \mathcal{D}$ domain of interest.

## 4.1 Linear Interpolation

In the two-dimensional case, bilinear interpolation is given by

$$\hat{L}(x, y) = \frac{1}{(x_2 - x_1)(y_2 - y_1)} \begin{bmatrix} x_2 - x & x - x_1 \end{bmatrix} \begin{bmatrix} \hat{L}(x_1, y_1) & \hat{L}(x_1, y_2) \\ \hat{L}(x_2, y_1) & \hat{L}(x_2, y_2) \end{bmatrix} \begin{bmatrix} y_2 - y \\ y - y_1 \end{bmatrix} \tag{27}$$

where $\{(x_i, y_j) : i, j = 1, 2\}$ are the four closest points to $(x, y)$ such that $x_1 \leq x \leq x_2$ and $y_1 \leq y \leq y_2$. This result can be derived by performing one-dimensional linear interpolation twice as shown in [1].

An alternative way to write this result will shed more lights in this context as it relates closer to regression techniques that we will introduce below. From the definition of linear interpolation, one shall realize that $\hat{L}(x, y)$ only depends on $x$, $y$, and $xy$. Therefore, we can write

$$\hat{L}(x, y) \approx \alpha_0 + \alpha_1 x + \alpha_2 y + \alpha_3 xy \tag{28}$$

where coefficients $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ are locally constant within each rectangle bounded by the closest points $\{(x_i, y_j) : i, j = 1, 2\}$ to $(x, y)$ as mentioned above. Moreover, by definition of linear interpolation, we know that $\hat{L}(x_i, y_j)$ fits exactly on the closest points. Thus, the coefficients can be determined by solving the linear system

$$\begin{bmatrix} \hat{L}(x_1, y_1) \\ \hat{L}(x_1, y_2) \\ \hat{L}(x_2, y_1) \\ \hat{L}(x_2, y_2) \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 & x_1 y_1 \\ 1 & x_1 & y_2 & x_1 y_2 \\ 1 & x_2 & y_1 & x_2 y_1 \\ 1 & x_2 & y_2 & x_2 y_2 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \tag{29}$$

With solved coefficients $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$, any $\hat{L}(x, y)$ within the small neighborhood can be determined.

## 4.2 Least-Square Monte Carlo

The idea of *least-square Monte Carlo* is to expand feature space based on underlying risk factors and approximate the function on the space spanned by expanded features. Let $\mathbf{X} \in \mathbb{R}^{N\times K}$ denotes the matrix of raw risk factors where each row $\mathbf{x}_i = (x_i^{(k)})_{k=1}^K$ is a scenario of $K$ risk factors. Feature expansion is a function $\Phi : \mathbb{R}^K \to \mathbb{R}^D$ acts on each row of $\mathbf{X}$ which produces expanded features $\Phi(\mathbf{X})$. For example, in the linear interpolation section, $\Phi : (x, y) \mapsto (1, x, y, xy)$. Equipped with expanded features, we can approximate our function as

$$\hat{L}(x, y) \approx \boldsymbol{\beta}^\top \Phi(x, y) \tag{30}$$

where $\boldsymbol{\beta} \in \mathbb{R}^D$ is a column vector of coefficients to be determined.

In practice, a popular choice is use polynomial expansion for $\Phi$ and minimize the sum of squared errors, collectively known as *polynomial regression.*

$$\arg\min_{\boldsymbol{\beta}\in\mathbb{R}^D} \sum_{i=1}^{N} \left(\hat{L}(x_i, y_i) - \boldsymbol{\beta}^\top \Phi(x_i, y_i)\right)^2 \tag{31}$$

where $\Phi : (x, y) \mapsto (x^u y^v)_{u+v\le p}$ produces a polynomial of degree $p$. With an abuse of notation, we overload $\mathbf{X}$ to denote the expanded features from now on. In statistics, it is commonly used to as the *design matrix.* From linear algebra, we know that if the number of distinct rows of $\mathbf{X}$ is no less than the dimension of $\boldsymbol{\beta}$, then the minimizer exists and is uniquely given by

$$\widehat{\boldsymbol{\beta}}^{\text{LS}} = (\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{Y} \tag{32}$$

where $\mathbf{Y} = \left(\hat{L}(x_i, y_i), \cdots, \hat{L}(x_n, y_n)\right)$. Once an estimate $\widehat{\boldsymbol{\beta}}$ is determined, we can easily approximate $\hat{L}$.

The choices of polynomials as basis function $\Phi$ and least-square minimization setup may seem somewhat arbitrary. We will try to provide some insights in the following sections.

### 4.2.1 Functional Approximation via Taylor Expansion

In this subsection, we aim to justify the usage of polynomial feature expansion. First, we state the Taylor's expansion at $(x_0, y_0)$ of any function $f$ with two independent variables $(x, y)$,

$$f(x_0 + h, y_0 + k) = \sum_{n=0}^{p} \frac{1}{n!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f(x_0, y_0) + R_p \tag{33}$$

where Lagrange remainder $R_N$ is given by

$$R_p = \frac{1}{(p+1)!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^{p+1} f(x_0 + \theta h, y_0 + \theta h) \tag{34}$$

with some $\theta \in [0, 1]$. From analysis, we know that Taylor's theorem holds (i.e. we can always find a $\theta \in [0, 1]$) for any combinations of $x_0$, $y_0$, $h$, $k$, and $N$ as long as all of the involves derivatives exist.

The exact expression of reminder is not very useful here as it requires too many information and does not provide analytic solution for $\theta$. Therefore, we will find an upper bound instead. Let $M$ denotes the supremum of $f$ around a neighborhood of $(x_0, y_0)$, and $\rho := \sqrt{h^2 + k^2}$, with Stirling's approximation for factorials, we can get

$$|R_p| \leq \frac{M}{\sqrt{2\pi(p+1)}}\left(\frac{\sqrt{2}\rho e}{p}\right)^{p+1} \tag{35}$$

With this approximate upper bound, we can establish a theoretical bound for any order of polynomial approximation if given information of upper bound of $f$, input data, and estimation region. Conversely, we can also determine the order of polynomial $p$ needed to achieve any prescribed level of accuracy $|R_p|$.

Connecting back to the previous examples, we work out the case of second-order polynomial expansion. To see this, we apply second order Taylor's expansion and substitute with $h = x - x_0$ and $k = y - y_0$

$$\begin{aligned} f(x, y) =& f(x_0, y_0) \\ &+ (x - x_0)\frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0)\frac{\partial f}{\partial y}(x_0, y_0) \\ &+ \frac{1}{2}(x - x_0)^2\frac{\partial^2 f}{\partial x^2}(x_0, y_0) + \frac{1}{2}(y - y_0)^2\frac{\partial^2 f}{\partial y^2}(x_0, y_0) + (x - x_0)(y - y_0)\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \\ &+ R_2 \\ =& Ax^2 + Bxy + Cy^2 + Dx + Ey + F + R_2 \end{aligned}$$

where $A$, $B$, $C$, $D$, $E$, $F$ are constants depending on $x_0$ and $y_0$. These constants can be identified from the equation,

$$
\begin{aligned}
A &= \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x_0, y_0) \quad B = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \quad C = \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(x_0, y_0) \\
D &= \frac{\partial f}{\partial x}(x_0, y_0) - x_0 \frac{\partial^2 f}{\partial x^2}(x_0, y_0) - y_0 \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \\
E &= \frac{\partial f}{\partial y}(x_0, y_0) - y_0 \frac{\partial^2 f}{\partial y^2}(x_0, y_0) - x_0 \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \\
F &= -x_0 \frac{\partial f}{\partial x}(x_0, y_0) - y_0 \frac{\partial f}{\partial y}(x_0, y_0) + \frac{1}{2}{x_0}^2 \frac{\partial^2 f}{\partial x^2}(x_0, y_0) + \frac{1}{2}{y_0}^2 \frac{\partial^2 f}{\partial y^2}(x_0, y_0) + x_0 y_0 \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)
\end{aligned}
$$

Based on the above results, the error term $R_2$ is bounded above in absolute value by some other constant depending on $x_0$, $y_0$, $M$, and $\sup\{(x - x_0)^2 + (y - y_0)^2\}$. Notice that this theoretical bound is contingent on knowing the values of derivatives composing coefficients $A$ through $F$, which is not the case in practice. The estimation of such coefficients will inevitably introduce additional errors. We will address this issue in the next subsection.

#### 4.2.2 Statistical Interpretation of Least-Square Parameter Estimation

In this subsection, we will provide a statistical perspective on least-square estimation. *Maximum likelihood estimation* is arguably the most widely-used parameter estimation strategy. As its names suggests, it maximizes the likelihood of observing the given data over the parameter space. Coincidentally, the least-square problem minimizes some quantity depending on the underlying parameters. To bridge the gap, we hope to connect these two expressions.

Consider the linear regression model $\mathbf{Y} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta}, \sigma^2 I_N)$ where response $\mathbf{Y} \in \mathbb{R}^N$, design matrix $\mathbf{X} \in \mathbb{R}^{N \times D}$, and coefficients $\boldsymbol{\beta} \in \mathbb{R}^D$. Its likelihood is given by Gaussian density

$$\mathcal{L}(\boldsymbol{\beta}, \sigma^2; \mathbf{Y}, \mathbf{X}) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left(-\frac{||\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}||_2^2}{2\sigma^2}\right)$$

It can be shown that maximize this quantity $\mathcal{L}$ is equivalent to

$$\min ||\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}||_2^2 + n\sigma^2 \ln \sigma^2$$

Despite potentially having an additional parameter $\sigma^2$ to maximize over, we have arrived at the same expression as the least-square minimization problem. This correspondence tells us that the least-square formulation is implicitly assuming an underlying normal model on approximation errors $\boldsymbol{\epsilon} := \mathbf{Y} - \mathbf{X}\boldsymbol{\beta}$. Furthermore, the errors shall be independently and identically distributed, i.e. $\epsilon_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$ for $\forall i \in [N]$. Inspired by this interpretation, one apply diagnosis in regression modeling to assess how well the assumptions are met, which is equivalent to argue how well the least-square estimation performs.

One common quantity to inspect is the variance of the least-square estimator, which is given by

$$\text{Var}[\widehat{\boldsymbol{\beta}}^{\text{LS}}] = \sigma^2(\mathbf{X}^\top \mathbf{X})^{-1}$$

It quantifies how precise our estimate of coefficients are given data $\mathbf{X}$. To get a better estimate, we may want *design* a good matrix $\mathbf{X}$. (Side note: this is why $\mathbf{X}$ is called the design matrix.) As suggests by the formula, we should sample mesh points such that $(\mathbf{X}^\top \mathbf{X})^{-1}$ is small in some sense (e.g., Frobenius norm).

# 5 Conclusion and Final Remarks

In this primer, we introduced/reviewed concepts surrounding GMMB contracts as an example to study equity-linked insurance products. In Section 2, we described how the contact operates and introduced pricing framework via Black-Scholes PDE and risk-neutral pricing; In Section 3 we further detailed hedging techniques including continuously hedging with exogenous cash flows and greeks letters hedging. In Section 4, we analyzed interpolation techniques and provided interpretation to some computational choices.

We hope this primer can inspire interests of beginners and review important framework for more advanced readers. Details of Section 2 and Section 3 lead to further study in mathematical finance. Interpolation topics in Section 4 are covered in numerical analysis. The last part on statistics is generally discussed in regression design/regression modeling. A visualization tool to demonstrate many concepts covered in this primer is under development. Please contact Feng [1] or Cui [2] for access to the currently developing version.

[1] Runhuan Feng, Director of Actuarial Science, rfeng@illinois.edu

[2] Haoen Cui, B.S. in Mathematics, Statistics, Finance, and Accountancy (2018), hcui10@illinois.edu

## References

[1] Runhuan Feng. *An Introduction to Computational Risk Management of Equity-Linked Insurance*. CRC Press, 1st edition, 2018 (Forthcoming).