APPLICATIONS OF MATHEMATICS IN FINANCE

U. N. KNOWN

- 1. Introduction. The topic this paper will cover is Applications of Mathematics in Finance. Although the paper will focus primarily on Linear Algebra and Calculus, the topic is predominately concerned with probability and statistics, as the goal of Financial Modeling is to predict the value of assets in the future, given the history of values. The most historic topic, Brownian Motion, is named in honor of a botanist Robert Brown in 1827, however the math behind it was not formalized by Albert Einstein until 1905; the finance application is using Brownian Motion to model asset value. After this discovery, we explore a different investment strategy published by Harry Markowitz in 1952. Instead of a continuous time approach, Markowitz attempts to minimize volatility in a stock portfolio with a covariance matrix. The miscellaneous topic is a hybrid of the first two topics. We consider the problem of filtering additive white noise from an asset's time series representation; we do this (similarly to the second topic), by considering the series's correlation matrix and demonstrate how this helps to filter the noise.
- 2. Advanced Calculus Math Concept. In Module Four, existence of functions that were continuous everywhere and differentiable nowhere was briefly discussed. Brownian motion is an example of such a function that models diffusion processes. Standard Brownian Motion obeys four properties: (1) W(0) = 0, (2) For $0 < t < s < u, t, s, u \in \mathbb{R}$, W(u - s) and W(s - t) are independent, (3) For $0 < t < s, t, s \in \mathbb{R}, W(t-s) \sim \mathcal{N}(0, t-s), \text{ and (4) For } t \in \mathbb{R} > 0, W(t) \text{ is continuous}$ in t. If W(0) is a different constant, then it is Non-Standard Brownian Motion. This is the most common continuous time process as its distribution comes from The Central Limit Theorem (third property), and it is independent of itself across time (second property). If we want to show that Brownian Motion has no hope of differentiability, we can take $\lim_{h\to 0} \frac{W(t+h)-W(t)}{k}$. Since W(t+h) and W(t) are independent, and each increment is $N(0,\Delta t)$, we get $W(t+h)-W(t)\sim \mathcal{N}(0,h)$. If we estimate the top with $\kappa\sqrt{h}, \kappa\in R$, we get $\lim_{h\to 0} \frac{\kappa\sqrt{h}}{h} = \lim_{h\to 0} \frac{\kappa}{\sqrt{h}} \to \infty$. This suggests that the jumps in Brownian Motion are too big to be differentiable. This can rigorously proven by using the Time Inversion Property of Brownian Motion (Corollary 1.11 in [2]). Stochastic Calculus can be derived from Brownian Motion by recognizing that the quadratic variation of Brownian Motion is Δt . The rules for Stochastic Calculus are as follows: $dt \times dt = 0$, $dt \times dW = 0$, and $dW \times dW = dt$. With Stochastic Calculus we can represent any $It\hat{o}$ Process as: $X_t = a_t dt + b_t dW$ [1].

I am interested in this application because it can take almost any continuous time process and represent it as a summation of a deterministic part and a random part. Intuitively, every process is either deterministic, random, or a combination. This is useful in finance because it models continuous assets very well. In order to truly appreciate this application, a more rigorous study of the $It\hat{o}$ integral is needed.

3. Linear Algebra Math Concept. Conceptually, the goal in managing a portfolio is to make the highest return possible at the lowest possible risk level; the Markowitz Portfolio is an optimization problem that allows one to achieve this under certain assumptions. Given a vector of stock returns in a portfolio \mathbf{x} , and the proportion of each stock we own \mathbf{w} , the total return $r = \mathbf{x}^T \mathbf{w}$. We can express the variance of r as $\sigma^2 = Cov(\mathbf{x}^T \mathbf{w}, \mathbf{x}^T \mathbf{w}) = \mathbf{w}^T Cov(\mathbf{x}, \mathbf{x}) \mathbf{w} = \mathbf{w}^T \Sigma \mathbf{w}$, where Σ is the

covariance matrix of \mathbf{x} . Now, the optimization problem is minimizing the variance of r under the constraint $\mathbf{w1} = 1$ (1 is vector of all ones; the weights must add up to one). Using Lagrange Multipliers, the optimal weight vector to minimize variance is $\mathbf{w}_{opt} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}}$ [4]. The linear algebra interpretation of this formula is that if $\Sigma = Q\Lambda Q^{-1}$ then $\Sigma^{-1} = Q^{-1}\Lambda^{-1}Q$. The top of the expression: $\Sigma^{-1}\mathbf{1}$ is a sum of the eigenvectors of Σ that gets scaled by the inverse eigenvalues, then laid back by Q^{-1} . The final expression is going to be dominated by the smallest eigenvalue of Σ . The bottom expression $\mathbf{w}^T\Sigma\mathbf{w}$ can be recognized as a Rayleigh Quotient, which normalizes the numerator. Therefore, we conclude that the optimal weight vector to minimize variance, very closely resembles $Q^{-1}\mathbf{v}$, where \mathbf{v} is the unit eigenvector that comes from the smallest eigenvalue of Σ . In the case where the smallest eigenvalue of Σ approaches zero, the resultant vector, $Q^{-1}\mathbf{v}$, becomes a one-hot vector, and we would conclude that one stock in our portfolio is much less volatile than the others.

I find this topic especially interesting, because the exact same optimization problem is used in Adaptive Beamforming. In the Engineering context, one focuses on the minimum variance response as minimum variance corresponds to the best signal integrity. In this context the weight vectors \mathbf{w} , actually corresponds to a spatial direction. In order to truly understand this application, it is important to understand how to take expectations of matrices.

4. Miscellaneous Math Concept. The previous section involved optimizing a portfolio with a covariance matrix of the different asset returns averaged over time. This next concept will involve a correlation matrix of different times (not categories). Consider a discrete time-series x, that is correlated with itself, for example $x_n = \rho x_{n-1} + \epsilon_i$ where ϵ_i is white (known as AR(1) model). Now suppose we want to eliminate the noise ϵ . This is not straightforward as it is unclear what part of x comes from previous x versus ϵ . To solve this, we use the Karhunen - Loève Transform. Define $\Sigma_{n,m} = E(x_n x_{n+m})$. We get that the Σ matrix is Symmetric and Positive-Definite (and Toeplitz if stationary). This matrix has an eigen-decomposition $\Sigma = Q\Lambda Q^H$. The Karhunen - Loève transform takes a blocks of data \mathbf{x} , and returns $Q^H\mathbf{x}$. If we examine $\Sigma' = E(Q^H\mathbf{x}_n\mathbf{x}_{n+m}^TQ) = Q^H(Q\Lambda Q^H)Q = \Lambda$ [3]. Diagonal Λ implies that the output of this transform is uncorrelated with itself. Because the terms are uncorrelated, the smaller terms in the transform are most likely to come from the additive noise. To filter out the noise, we round small values down and take the Inverse Karhunen - Loève Transform.

I enjoy this topic as it can be viewed as a generalization of the Fourier Transform for stochastic processes. The $Karhunen-Lo\`eve$ Transform expresses a time-series with respect to an orthonormal basis that comes from its correlation matrix. The one problem with this concept in practice is if the time series deviates from stationarity, its correlation matrix is no longer valid; this means a new one needs to be estimated from incoming data.

REFERENCES

- [1] M. Haugh, A brief introduction to stochastic calculus, 2016.
- [2] P. MORTERS AND Y. PERES, Brownian Motion, Cambridge University Press, Cambridge, UK, 2010.
- [3] M. U. TORIN, O. YILMAZ, AND A. N. AKANSU, Financial Signal Processing and Machine Learning, John Wiley and Sons, West Sussex, UK, 2016.
- [4] J. Zhu, Markowitz portfolio theory.