Optimal Investment Strategies under Delayed Information: General Theory, Specific Model Analysis, and AR(1) Application

BSc Thesis

Author:
Tamás Terényi

Adviser:
Miklós Rásonyi
Department of Probability Theory and Statistics

Mathematics BSc

Eötvös Loránd University
Faculty of Science
Budapest, 2025

Abstract

This thesis investigates optimal investment strategies within the framework of exponential utility maximization. The work is presented in three main parts.

First, we explain a general theory for an investor facing a delay D in observing a price process $S = (S_k)_{k=0,\dots,n}$ whose increments are multivariate normal developed by Dolinsky and Zuk (2023). They establish the existence and uniqueness of an optimal trading strategy $\hat{\gamma}$. This result hinges on a unique decomposition of the inverse covariance matrix (precision matrix) Λ of price increments into $\Lambda = \hat{Q}^{-1} + \hat{\Gamma}$, where $\hat{Q}$ is a banded positive-definite matrix and $\hat{\Gamma}$ is a symmetric matrix with a specific zero-pattern related to the delay D. The proof involves a verification theorem based on a martingale measure approach and duality arguments.

Second, to illustrate the practicalities of determining the crucial precision matrix Λ, we analyze a specific price process $S_i = z_{i-1} + z_i$, where z_k are i.i.d. Gaussian variables. For the increments $X_i = S_i - S_{i-1}$ of this MA(1)-like price model, we explicitly compute their covariance matrix Σ. The core of this part is the derivation of a closed-form expression for every entry of the precision matrix $\Lambda = \Sigma^{-1}$, using block matrix inversion techniques and properties of tridiagonal "path" matrices. This part showcases the difficulties of coming to an explicit solution despite strong results which make it theoretically possible.

Third, we apply the general optimal investment framework to a discrete-time Autoregressive (AR(1)) model $X_t = (1 + \beta)\, X_{t-1} + \sigma\, \epsilon_t$ for the risky asset, specifically under conditions of no information delay ($D = 0$) and a zero initial state ($X_0 = 0$). By first deriving the explicit precision matrix for the AR(1) increments and then substituting it into the simplified optimal strategy formula (where expected increments $\mu_j = 0$), we demonstrate that the resulting strategy, $\hat{\gamma}_i = \frac{\beta}{\sigma^2}[1 - (T - i)\beta]X_{i-1}$, precisely matches the known solution found by Deák and Rásonyi (2015) for fully informed investors. This reconciliation serves as a validation of both results.

Contents

Chapter 1

Introduction

Optimal investment decisions are a cornerstone of financial mathematics and economics. Investors constantly seek strategies to maximize their expected utility of wealth, balancing risk and return. Classical models often assume instantaneous availability of information. However, in real-world markets, information delays are prevalent, stemming from various sources such as data processing times, reporting lags, or the inherent nature of certain economic indicators, which can be of particular effect in high-frequency trading for instance (Föllmer and Schied, 2016, Chapter 3). This thesis addresses the problem of optimal portfolio selection for an investor who faces such information delays.

The primary objective is to develop a framework for determining optimal trading strategies under exponential utility when market information is not immediately available. We consider a discrete-time financial market where the price increments of a risky asset are multivariate normally distributed. The investor's information about past prices is subject to a fixed delay D.

This work is structured into three main parts:

The first part of the thesis (Chapter 2) lays down the general theoretical groundwork based on Dolinsky and Zuk (2023). We formulate the problem of maximizing expected exponential utility of terminal wealth in the presence of information delays. We prove a theorem that characterizes the unique optimal trading strategy. This characterization relies on a specific decomposition of the precision matrix. The proof of this theorem employs techniques from convex optimization and martingale duality.

The second part (Chapter 3) focuses on a use case of the general theory. To provide a concrete and non-trivial example of how the necessary matrix can be derived, we analyze a specific price process model where prices are formed by the sum of two consecutive i.i.d. Gaussian random variables. For this model, we explicitly calculate the covariance matrix of price increments and then undertake the derivation of its inverse. This involves leveraging results for the inverse of tridiagonal matrices and applying block matrix inversion techniques (Schur complements) to a reordered system.

The third part (Chapter 4) serves to connect the general theory with an established results in financial modeling by Deák and Rásonyi (2015). We consider a standard Autoregressive process of order 1 for the risky asset. Under the simplifying assumptions of no information delay ($D = 0$) and a zero initial asset value, we apply the optimal strategy formula derived in Chapter 2. We explicitly compute the precision matrix for the AR(1) increments and demonstrate that the general theory recovers the known optimal trading strategy for a fully informed investor in this AR(1) setting.

Chapter 2

Optimal Investment with Delayed Information

This chapter develops the general mathematical framework for determining the optimal investment strategy for an investor who maximizes exponential utility of terminal wealth while facing a delay in the observation of asset prices based on Dolinsky and Zuk (2023). We begin by defining the market model and the investor's optimization problem as they did, then explain their proof step by step.

2.1 Model Setup and Problem Formulation

We have a price process $S = (S_k)_{k=0,\ldots,n}$ on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The increments $(S_1 - S_0, \ldots, S_n - S_{n-1})$ are multivariate normal with mean vector μ and covariance matrix Σ, which is positive-definite. Set $\Lambda := \Sigma^{-1}$.

We fix a delay $D \in \mathbb{Z}_+$. The information filtration of our investor is $\mathcal{G}_k^D := \sigma(S_0, \ldots, S_{(k-D)^+})$. A trading strategy $\gamma = (\gamma_1, \ldots, \gamma_n)$ is predictable with respect to $\{\mathcal{G}_k^D\}$. Its terminal wealth at time n is

$$V_n^\gamma = \sum_{i=1}^{n} \gamma_i (S_i - S_{i-1})$$

We consider exponential utility with parameter 1 (i.e. risk aversion $\alpha = 1$), and the goal is:

$$\text{Maximize} \quad \mathbb{E}\left[-e^{-V_n^\gamma}\right] \quad \text{over all } \gamma \in \mathcal{A}_D$$

We introduce two classes of matrices:

1. $\mathcal{S}_D \subset M_n(\mathbb{R})$: the set of positive-definite matrices Q with banded structure such that $Q_{ij} = 0$ whenever $|i - j| > D$. (Hence Q is a covariance-type matrix with a bandwidth D.)

2. $\mathcal{T}_D \subset M_n(\mathbb{R})$: the set of symmetric matrices Γ such that $\Gamma_{ij} = 0$ whenever $|i - j| \leq D$. (Hence Γ has zeros on its main diagonal and on D diagonals on either side.)

2.2 Main Theorem on Optimal Strategy

The central result of this chapter provides the unique optimal trading strategy and the corresponding maximum expected utility. It relies on a specific algebraic decomposition of the precision matrix Λ. The theorem states that there exists a unique decomposition

$$\Lambda = \hat{Q}^{-1} + \hat{\Gamma} \tag{2.1}$$

where $\hat{Q} \in \mathcal{S}_D$ and $\hat{\Gamma} \in \mathcal{T}_D$. The maximizer $\hat{\gamma} = (\hat{\gamma}_1, ..., \hat{\gamma}_n)$ for the optimization problem is unique and is given by the linear form

$$\hat{\gamma}_i = \sum_{j=1}^{n} \Lambda_{ij}\mu_j - \sum_{j=1}^{i-1} \hat{\Gamma}_{ij}(S_j - S_{j-1}), \quad i = 1, ..., n. \tag{2.2}$$

The corresponding value is given by

$$\mathbb{E}_{\mathbb{P}}\left[-\exp(-V_n^{\hat{\gamma}})\right] = -\sqrt{\left|\frac{\hat{Q}}{\Sigma}\right|}\exp\left(-\frac{1}{2}\mu\Lambda\mu'\right) \tag{2.3}$$

2.3 Proof of the Main Theorem

The proof of this theorem proceeds in three main steps. First, we establish the existence and uniqueness of the matrix decomposition $\Lambda = \hat{Q}^{-1} + \hat{\Gamma}$, which we will not detail. Second, we present a verification argument based on martingale duality, which provides conditions for a candidate strategy to be optimal. Third, we construct the candidate optimal strategy $\hat{\gamma}$ and an associated dual measure $\hat{\mathbb{Q}}$ that satisfy these conditions.

2.3.1 Existence and Uniqueness of the Matrix Decomposition

This step demonstrates that the precision matrix Λ can be uniquely decomposed as stated in the theorem.

1. **Positivity and Sylvester's criterion:** We know Λ is positive-definite. Hence, all its leading principal minors are positive as per (Gilbert, 1991). In particular, the relevant minors which overlap only in certain banded parts are also positive.

2. **Application of Theorem 5.5 in Barrett and Feinsilver (1981):** Because Λ satisfies those positivity properties on principal minors (specifically the ones with "width" $D+1$), it follows from that result that Λ can be written as $\Lambda = R + (\Lambda - R)$, where R itself is invertible and banded (and in fact symmetric), and $\Lambda - R$ is in $\mathcal{T}_D$.

3. **Set** $\hat{Q} := R^{-1}$. Then $\hat{Q}$ is in $\mathcal{S}_D$ because it's the inverse of a banded, positive-definite matrix R. (Equivalently, the structure in $\hat{Q}$ and positivity follow from the structure and positivity of R.) We define $\hat{\Gamma} := \Lambda - R$. By construction, $\hat{\Gamma}$ is in $\mathcal{T}_D$. And we get

$$\Lambda = \hat{Q}^{-1} + \hat{\Gamma}.$$

4. **Uniqueness:** The same reference ensures that if you try to represent Λ in that form in two different ways, they must coincide.

So Step I concludes that there is a unique decomposition

$$\Lambda = \hat{Q}^{-1} + \hat{\Gamma}, \quad \hat{Q} \in \mathcal{S}_D, \ \hat{\Gamma} \in \mathcal{T}_D$$

2.3.2 Verification via Martingale Duality

This step introduces a verification result. If we can find a strategy $\tilde{\gamma}$, an equivalent probability measure $\tilde{\mathbb{Q}}$, and a constant C satisfying a specific relationship, then $\tilde{\gamma}$ is guaranteed to be the optimal strategy.

The Setup for Step II

We define $\mathcal{Q}_D$ to be the collection of probability measures $\mathbb{Q}$ (equivalent to $\mathbb{P}$) that have finite relative entropy and also satisfy a certain conditional mean-zero condition, namely:

$$\mathbb{E}_{\mathbb{Q}}[S_t - S_s \mid \mathcal{G}_s^D] = 0 \quad \text{for all } t \geq s. \tag{2.1}$$

This effectively says that under $\mathbb{Q}$, the increments $S_t - S_s$ behave like martingale increments with respect to the delayed filtration $\{\mathcal{G}_s^D\}$.

The Statement in Step II

Step II says:

> If we manage to find a triplet $(\tilde{\gamma}, \tilde{\mathbb{Q}}, C)$ with $\tilde{\gamma} \in \mathcal{A}_D$, $\tilde{\mathbb{Q}} \in \mathcal{Q}_D$, and a real constant C such that
>
> $$V_n^{\tilde{\gamma}} + \log\left(\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right) = C \quad \text{(almost surely)}, \tag{2.2}$$
>
> then $\tilde{\gamma}$ must be the unique optimal portfolio for the exponential utility objective, and
>
> $$\mathbb{E}_{\mathbb{P}}\left[-e^{-V_n^{\tilde{\gamma}}}\right] = -e^{-C}$$

So Step II is a verification step: if you guess (or derive) the right strategy and the right measure $\tilde{\mathbb{Q}}$ so that (2.2) is fulfilled, that guess is forced to be correct.

Why This Implies Optimality

1. **Relation (2.2) implies**

   $$\mathbb{E}_{\mathbb{P}}\left[-e^{-V_n^{\tilde{\gamma}}}\right] = -e^{-C}.$$

2. **Uniqueness:** For exponential utility $-e^{-x}$, the maximization problem is $\max_\gamma \mathbb{E}\left[-e^{-V_n^\gamma}\right]$. This is a strictly concave functional in γ. Hence, if an optimizer exists, it is unique. So any candidate that we can verify is an optimizer must automatically be the unique one.

3. **Verifying that no other γ can yield a strictly larger value** is done via a standard "duality" or "Fenchel–Legendre" inequality argument:

   $$\log\left(\mathbb{E}_{\mathbb{P}}\left[e^{-V_n^\gamma}\right]\right) \geq -\mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] \tag{2.4}$$

   for any $\mathbb{Q} \in \mathcal{Q}_D$.

Why (2.1) ensures $\mathbb{E}_{\mathbb{Q}}[V_n^\gamma] = 0$

Observe that condition (2.1) says $\mathbb{E}_{\mathbb{Q}}[S_t - S_s \mid \mathcal{G}_s^D] = 0$. For a delayed-predictable strategy γ, you effectively can factor out γ_i as a $\mathcal{G}_{i-1}^D$-measurable random variable in the difference $S_i - S_{i-1}$. Then a repeated application of (2.1) yields

$$\mathbb{E}_{\mathbb{Q}}[V_n^\gamma] = \mathbb{E}_{\mathbb{Q}}\left[\sum_{i=1}^n \gamma_i(S_i - S_{i-1})\right] = 0$$

Thus under $\mathbb{Q} \in \mathcal{Q}_D$, the expected increment of the wealth process is zero.

Derivation of the inequality (2.3) from (2.2)

- **(2.2)** says that for some strategy $\tilde{\gamma} \in \mathcal{A}_D$, measure $\tilde{\mathbb{Q}} \in \mathcal{Q}_D$, and constant $C \in \mathbb{R}$,

$$V_n^{\tilde{\gamma}} + \log\left(\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right) = C \quad \text{(almost surely).}$$

- we are aiming for

$$\log\left(\mathbb{E}_{\mathbb{P}}\left[e^{-V_n^{\tilde{\gamma}}}\right]\right) = -C = -\mathbb{E}_{\tilde{\mathbb{Q}}}\left[\log\left(\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right)\right] \tag{2.3}$$

as in the original proof. We want to see why both equalities in (2.3) hold.

From (2.2), we have

$$V_n^{\tilde{\gamma}} + \log\left(\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right) = C$$

Rearrange to

$$-V_n^{\tilde{\gamma}} = -C + \log\left(\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right)$$

Exponentiate both sides to get

$$e^{-V_n^{\tilde{\gamma}}} = e^{-C} \exp\left(\log\left(\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right)\right) = e^{-C}\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}$$

Taking expectation under $\mathbb{P}$:

$$\mathbb{E}_{\mathbb{P}}\left[e^{-V_n^{\tilde{\gamma}}}\right] = e^{-C}\mathbb{E}_{\mathbb{P}}\left[\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right]$$

But by definition of the Radon–Nikodým derivative

$$\mathbb{E}_{\mathbb{P}}\left[\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right] = \tilde{\mathbb{Q}}\left(\Omega\right) = 1$$

So

$$\mathbb{E}_{\mathbb{P}}\left[e^{-V_n^{\tilde{\gamma}}}\right] = e^{-C}$$

Taking the natural logarithm gives

$$\log\left(\mathbb{E}_{\mathbb{P}}\left[e^{-V_n^{\tilde{\gamma}}}\right]\right) = \log(e^{-C}) = -C$$

Again from (2.2):

$$\log\left(\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right) = C - V_n^{\tilde{\gamma}}$$

Take expectation now under $\tilde{\mathbb{Q}}$:

$$\mathbb{E}_{\tilde{\mathbb{Q}}}\left[\log\left(\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right)\right] = \mathbb{E}_{\tilde{\mathbb{Q}}}[C] - \mathbb{E}_{\tilde{\mathbb{Q}}}[V_n^{\tilde{\gamma}}]$$

By definition, C is just a constant, so $\mathbb{E}_{\tilde{\mathbb{Q}}}[C] = C$. Also, $\tilde{\mathbb{Q}} \in \mathcal{Q}_D$ ensures $\mathbb{E}_{\tilde{\mathbb{Q}}}[V_n^{\tilde{\gamma}}] = 0$. Therefore,

$$\mathbb{E}_{\tilde{\mathbb{Q}}}\left[\log\left(\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right)\right] = C$$

Hence

$$-C = -\mathbb{E}_{\tilde{\mathbb{Q}}}\left[\log\left(\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right)\right]$$

Putting both pieces together yields exactly the chain of equalities in (2.3):

$$\log\left(\mathbb{E}_{\mathbb{P}}\left[e^{-V_n^{\tilde{\gamma}}}\right]\right) = -C = -\mathbb{E}_{\tilde{\mathbb{Q}}}\left[\log\left(\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}}\right)\right]$$

Derivation of the Duality Inequality

1. **We are trying to show** that for any strategy γ,

$$\log\left(\mathbb{E}_{\mathbb{P}}\left[e^{-V_n^{\gamma}}\right]\right) \geq -\mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right].$$

   This is a form of weak duality statement: the left side (the "primal" expression) is bounded below by the right side (the "dual" expression).

2. **Why it's "$\geq$":**
   - We want to minimize $\mathbb{E}[e^{-V}]$ (equivalently maximize $-\mathbb{E}[e^{-V}]$).
   - When you take the logarithm, minimizing $\mathbb{E}[e^{-V}]$ is the same as minimizing $\log(\mathbb{E}[e^{-V}])$.
   - You exhibit a particular strategy $\hat{\gamma}$ (the "candidate optimal strategy") and a particular measure $\hat{\mathbb{Q}}$ that satisfy this inequality and achieve equality.

Fix any $\gamma \in \mathcal{A}_D$. We want:

$$\log\left(\mathbb{E}_{\mathbb{P}}[e^{-V_n^{\gamma}}]\right) \geq -\mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right], \quad \text{for every } \mathbb{Q} \in \mathcal{Q}_D$$

Assume $\mathbb{E}_{\mathbb{P}}[e^{-V_n^{\gamma}}] < \infty$. Otherwise, the left-hand side is $+\infty$ and the inequality is trivial. Write

$$\mathbb{E}_{\mathbb{P}}[e^{-V_n^{\gamma}}] = \mathbb{E}_{\mathbb{P}}\left[e^{-V_n^{\gamma}} + z\frac{d\mathbb{Q}}{d\mathbb{P}}V_n^{\gamma}\right],$$

if and only if $\mathbb{E}_{\mathbb{Q}}[V_n^{\gamma}] = 0$. Since $\gamma \in \mathcal{A}_D$ and $\mathbb{Q} \in \mathcal{Q}_D$, we do have $\mathbb{E}_{\mathbb{Q}}[V_n^{\gamma}] = 0$, so we can add that zero term. Now apply the Legendre–Fenchel (or Young) inequality:

$$xy \leq e^x + y(\log y - 1),$$

valid for all real x, y. In our proof, we set $x = -V_n^\gamma(\omega)$ and $y = z\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)$. Then:

$$-V_n^\gamma(\omega) \cdot \left(z\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)\right) \leq e^{-V_n^\gamma(\omega)} + z\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)\left(\log\left(z\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)\right) - 1\right)$$

Add $z\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)V_n^\gamma(\omega)$ to both sides. The left becomes 0:

$$0 \leq \left[e^{-V_n^\gamma(\omega)} + z\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)\left(\log(\cdots) - 1\right)\right] + z\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)V_n^\gamma(\omega)$$

Regroup:

$$\underbrace{\left[e^{-V_n^\gamma(\omega)} + z\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)V_n^\gamma(\omega)\right]}_{\text{call this "X"}} + \underbrace{z\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)\left(\log(\cdots) - 1\right)}_{\text{call this "Y"}}$$

So:

$$0 \leq X + Y \Rightarrow X \geq -Y$$

Rewriting gives:

$$X \geq -Y = z\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)\left(1 - \log\left(z\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)\right)\right)$$

Substitute back:

$$e^{-V_n^\gamma(\omega)} + z\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)V_n^\gamma(\omega) \geq z\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)\left(1 - \log\left(z\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)\right)\right)$$

Taking expectations under $\mathbb{P}$:

$$\mathbb{E}_{\mathbb{P}}\left[e^{-V_n^\gamma} + z\frac{d\mathbb{Q}}{d\mathbb{P}}V_n^\gamma\right] \geq \mathbb{E}_{\mathbb{P}}\left[z\frac{d\mathbb{Q}}{d\mathbb{P}}\left(1 - \log\left(z\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right)\right]$$

Since z is constant:

$$\mathbb{E}_{\mathbb{P}}\left[z\frac{d\mathbb{Q}}{d\mathbb{P}}\left(1 - \log\left(z\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right)\right] = z\mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\left(1 - \log\left(z\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right)\right]$$

Distribute inside the expectation:

$$= z\left(\mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\right] - \mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\log\left(z\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]\right)$$

Since $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is the Radon–Nikodym derivative:

$$\mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\right] = \mathbb{Q}(\Omega) = 1$$

For the second term:

$$\mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\log\left(z\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] = \mathbb{E}_{\mathbb{Q}}\left[\log\left(z\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]$$

Thus:

$$z\left(1 - \mathbb{E}_{\mathbb{Q}}\left[\log\left(z\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]\right)$$

Expand the logarithm:

$$\mathbb{E}_{\mathbb{Q}}\left[\log\left(z\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] = \log z + \mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]$$

Therefore:

$$1 - \mathbb{E}_{\mathbb{Q}}\left[\log\left(z\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] = 1 - \log z - \mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right],$$

and multiplying by z:

$$z - z\log z - z\mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]$$

Putting everything together:

$$\mathbb{E}_{\mathbb{P}}\left[e^{-V_n^\gamma}\right] \geq z - z\log z - z\mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]$$

Taking logarithms gives the desired inequality.

Define

$$f(z) := z - z\log z - z\mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right], \quad z > 0$$

For convenience, let

$$\alpha := \mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]$$

Then

$$f(z) = z - z\log z - \alpha z = z\left[1 - \log z - \alpha\right]$$

Our goal is to find $\max_{z>0} f(z)$. Because $f(z)$ is differentiable on $(0,\infty)$ and tends to $-\infty$ as $z \to 0^+$ or $z \to \infty$, it will have a global maximum at a critical point where $f'(z) = 0$.

We compute $f'(z)$ by breaking f into simpler terms:

1. The derivative of z w.r.t. z is 1.
2. The derivative of $-z\log z$ is $-(\log z + 1)$.
3. The derivative of $-\alpha z$ is $-\alpha$.

Putting these together:

$$f'(z) = 1 - (\log z + 1) - \alpha = -\log z - \alpha$$

Hence

$$f'(z) = 0 \iff -\log z - \alpha = 0 \iff \log z = -\alpha \iff z = e^{-\alpha}$$

So the **unique critical point** is

$$z^* = e^{-\alpha} = \exp\left(-\mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]\right)$$

Because $f(z)$ is a concave function on $(0,\infty)$ (one can also check $f''(z) < 0$ for $z > 0$), this critical point is indeed the **global maximum**.

To see the maximum value, we plug z^* into f. That is,

$$f(z^*) = f(e^{-\alpha}) = e^{-\alpha} - (e^{-\alpha})\log(e^{-\alpha}) - \alpha e^{-\alpha}$$

Recall: $\log(e^{-\alpha}) = -\alpha$, so

$$(e^{-\alpha})\log(e^{-\alpha}) = e^{-\alpha}(-\alpha) = -\alpha e^{-\alpha}$$

Thus,

$$f(e^{-\alpha}) = e^{-\alpha} - (-\alpha e^{-\alpha}) - \alpha e^{-\alpha} = e^{-\alpha} + \alpha e^{-\alpha} - \alpha e^{-\alpha} = e^{-\alpha}$$

Hence the maximum value of $f(z)$ over $z > 0$ is

$$\max_{z>0} f(z) = e^{-\alpha} = \exp\left(-\mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]\right)$$

Therefore,

$$\mathbb{E}_{\mathbb{P}}\left[e^{-V_n^{\gamma}}\right] \geq \exp\left(1 - \mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]\right),$$

which is equivalent to

$$\log\left(\mathbb{E}_{\mathbb{P}}\left[e^{-V_n^{\gamma}}\right]\right) \geq -\mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]$$

2.3.3 Construction of the Candidate Optimal Strategy and Dual Measure

The final step in the proof is to explicitly construct the strategy $\hat{\gamma}$ (as given in the theorem statement) and a measure $\hat{\mathbb{Q}}$ such that they satisfy the verification condition (2.2) from Step II. This confirms that $\hat{\gamma}$ is the unique optimal strategy.

First, we explicitly define the candidate optimal trading strategy

$$\hat{\gamma}_i = \sum_{j=1}^{n} \Lambda_{ij}\mu_j - \sum_{j=1}^{i-1} \hat{\Gamma}_{ij}(S_j - S_{j-1}), \quad i = 1, \ldots, n,$$

and verify that $\hat{\gamma}$ is admissible, i.e., $\hat{\gamma} \in \mathcal{A}_D$. Second, we construct a probability measure $\hat{\mathbb{Q}}$ (via its Radon–Nikodym derivative with respect to $\mathbb{P}$) such that the equality

$$V_n^{\hat{\gamma}} + \log\left(\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}}\right) = C$$

holds, where the constant C is defined by

$$C = \frac{1}{2}\left(\log|\Sigma| - \log|\hat{Q}| + \mu\Lambda\mu'\right)$$

This identity is exactly the condition needed (as established in Step II) to verify that $\hat{\gamma}$ is the unique optimizer for our exponential utility maximization problem.

Sub-step 1: Verification that $\hat{\gamma}$ is Admissible

1. Definition of $\hat{\gamma}$: The candidate strategy is given by

$$\hat{\gamma}_i = \sum_{j=1}^{n} \Lambda_{ij}\mu_j - \sum_{j=1}^{i-1} \hat{\Gamma}_{ij}(S_j - S_{j-1}), \quad i = 1, \ldots, n$$

Recall that the matrices Λ and $\hat{\Gamma}$ arise from the unique decomposition

$$\Lambda = \hat{Q}^{-1} + \hat{\Gamma},$$

where $\hat{Q} \in \mathcal{S}_D$ (a banded, positive-definite matrix) and $\hat{\Gamma} \in \mathcal{T}_D$ (a symmetric matrix with $\hat{\Gamma}_{ij} = 0$ whenever $|i - j| \leq D$).

2. Admissibility Check: The strategy $\hat{\gamma}$ is adapted to the filtration $\{\mathcal{G}_k^D\}$ because the second term in the expression for $\hat{\gamma}_i$ involves summation over indices $j \leq i-1$. Moreover, given that $\hat{\Gamma}_{ij} = 0$ for $|i-j| \leq D$, the remaining non-zero terms in the sum involve only increments that are $\mathcal{G}_{i-1}^D$-measurable. Specifically, $S_j - S_{j-1}$ is $\mathcal{G}_j^D$-measurable. For $\hat{\gamma}_i$ to be $\mathcal{G}_{i-1}^D$-measurable, we need $S_j - S_{j-1}$ involved in the sum to be $\mathcal{G}_{i-1}^D$-measurable. This means $j \leq (i-1-D)^+$. However, the sum goes up to $j = i-1$. The condition $\hat{\Gamma}_{ij} = 0$ if $|i-j| \leq D$ means that for $\hat{\Gamma}_{i,j}$ to be non-zero, we need $|i-j| > D$. If $j < i$, this means $i - j > D$, or $j < i - D$. This ensures $S_j - S_{j-1}$ is $\mathcal{G}_{(i-D-1)^+}^D \subseteq \mathcal{G}_{i-1}^D$-measurable. Thus, $\hat{\gamma}$ is predictable with respect to the delayed information flow, and we conclude $\hat{\gamma} \in \mathcal{A}_D$.

Sub-step 2: Definition of the Constant C and the Dual Measure $\hat{\mathbb{Q}}$

1. Constant C:

Define

$$C = \frac{1}{2}\left(\log|\Sigma| - \log|\hat{Q}| + \mu\Lambda\mu'\right)$$

This constant is introduced in order to normalize the exponential of the terminal wealth, as will be seen below.

2. Definition of $\hat{\mathbb{Q}}$:

The measure $\hat{\mathbb{Q}}$ is defined via its Radon–Nikodym derivative with respect to $\mathbb{P}$ by

$$\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}} = \exp(C - V_n^{\hat{\gamma}})$$

Notice that if the identity

$$V_n^{\hat{\gamma}} + \log\left(\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}}\right) = C$$

holds pointwise, then taking expectations under $\mathbb{P}$ yields

$$\mathbb{E}_{\mathbb{P}}\left[\exp(-V_n^{\hat{\gamma}})\right] = \exp(-C),$$

which, via Step II, guarantees that $\hat{\gamma}$ is optimal. To prove that $\hat{\mathbb{Q}}$ is a probability measure (and to compute its effect on the distribution of the increments), we proceed as follows:

1. Expressing the Wealth Process in Terms of Increments:

Let $X_i = S_i - S_{i-1}$ for $i = 1, \ldots, n$, so that the terminal wealth becomes

$$V_n^{\hat{\gamma}} = \hat{\gamma}_1 X_1 + \cdots + \hat{\gamma}_n X_n = \hat{\gamma} X'$$

Recall that

$$V_n^{\hat{\gamma}} = \sum_{i=1}^n \hat{\gamma}_i (S_i - S_{i-1}) = \sum_{i=1}^n \hat{\gamma}_i X_i$$

Hence, if we view $\hat{\gamma}$ and X both as row vectors, $\hat{\gamma} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_n)$ and $X = (X_1, \ldots, X_n)$, then

$$V_n^{\hat{\gamma}} = \hat{\gamma} X' = \sum_{i=1}^n \hat{\gamma}_i X_i$$

By definition (from the proof's construction), each component $\hat{\gamma}_i$ of the optimal strategy takes the form

$$\hat{\gamma}_i = \sum_{j=1}^{n} \Lambda_{ij}\mu_j - \sum_{j=1}^{i-1} \hat{\Gamma}_{ij} X_j$$

Hence

$$V_n^{\hat{\gamma}} = \sum_{i=1}^{n} \left[\sum_{j=1}^{n} \Lambda_{ij}\mu_j - \sum_{j=1}^{i-1} \hat{\Gamma}_{ij} X_j \right] X_i$$

Rewrite $V_n^{\hat{\gamma}}$ by splitting it into "$\Lambda\mu$" terms and "$\hat{\Gamma}$" terms:

$$V_n^{\hat{\gamma}} = \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{n} \Lambda_{ij}\mu_j X_i}_{\text{Term (A)}} - \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{i-1} \hat{\Gamma}_{ij} X_j X_i}_{\text{Term (B)}}$$

We next handle each part separately.

$$(\text{A}) = \sum_{i=1}^{n}\sum_{j=1}^{n} \Lambda_{ij}\mu_j X_i$$

If one interprets X as a row vector and μ as a row vector, then in matrix notation this double sum is the same as the scalar $X(\Lambda\mu')$. Equivalently, one might write

$$(\text{A}) = \mu\Lambda X'$$

Both interpretations reflect that we are summing over the product of $\Lambda_{ij}\mu_j X_i$ in index form. Either way,

$$(\text{A}) = \mu\Lambda X'$$

$$(\text{B}) = \sum_{i=1}^{n}\sum_{j=1}^{i-1} \hat{\Gamma}_{ij} X_j X_i$$

Because $\hat{\Gamma}$ is symmetric, i.e. $\hat{\Gamma}_{ij} = \hat{\Gamma}_{ji}$, we can symmetrize the sum over $i \neq j$. Concretely,

$$\sum_{i=1}^{n}\sum_{j=1}^{i-1} \hat{\Gamma}_{ij} X_j X_i = \sum_{1 \leq j < i \leq n} \hat{\Gamma}_{ij} X_j X_i = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \hat{\Gamma}_{ij} X_i X_j$$

Hence, in matrix form,

$$(\text{B}) = \frac{1}{2} X\hat{\Gamma}X'$$

Putting (A) and (B) together, we obtain:

$$V_n^{\hat{\gamma}} = [(\text{A})] - [(\text{B})] = \mu\Lambda X' - \frac{1}{2} X\hat{\Gamma}X'$$

That is precisely the desired identity:

$$V_n^{\hat{\gamma}} = \mu\Lambda X' - \frac{1}{2} X\hat{\Gamma}X'$$

Recall that $\Lambda = \hat{Q}^{-1} + \hat{\Gamma}$. It then follows that

$$V_n^{\hat{\gamma}} = \mu\Lambda X' - \frac{1}{2} X\hat{\Gamma}X' = \mu\Lambda X' - \frac{1}{2} X(\Lambda - \hat{Q}^{-1})X' = \mu\Lambda X' - \frac{1}{2} X\Lambda X' + \frac{1}{2} X\hat{Q}^{-1}X'$$

The construction permits one to express the Radon–Nikodym derivative as a ratio of two Gaussian densities. Specifically, since under $\mathbb{P}$ the vector $X = (X_1, \ldots, X_n)$ is distributed as

$$X \sim \mathcal{N}(\mu, \Sigma),$$

A crucial step is the explicit formula for $V_n^{\hat{\gamma}}$ in terms of X. By construction (using the symmetry of $\hat{\Gamma}$ and the decomposition $\Lambda = \hat{Q}^{-1} + \hat{\Gamma}$), one shows:

$$V_n^{\hat{\gamma}} = \mu\Lambda X' - \frac{1}{2}X\hat{\Gamma}X'$$

But since $\hat{\Gamma} = \Lambda - \hat{Q}^{-1}$, the expression rearranges to

$$V_n^{\hat{\gamma}} = \underbrace{\mu\Lambda X' - \frac{1}{2}X\Lambda X'}_{\text{“shift by }\Lambda\text{"}} + \frac{1}{2}X\hat{Q}^{-1}X'$$

Hence

$$-V_n^{\hat{\gamma}} = -\mu\Lambda X' + \frac{1}{2}X\Lambda X' - \frac{1}{2}X\hat{Q}^{-1}X'$$

The constant C is chosen (by design) so that when you exponentiate

$$C - V_n^{\hat{\gamma}},$$

it lines up exactly with the ratio of two Gaussian likelihoods. Concretely, recall

$$C = \frac{1}{2}(\log|\Sigma| - \log|\hat{Q}| + \mu\Lambda\mu')$$

Combining everything, one writes

$$\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}} = \exp(C - V_n^{\hat{\gamma}}) = \exp\left(\frac{1}{2}(\log|\Sigma| - \log|\hat{Q}| + \mu\Lambda\mu') - V_n^{\hat{\gamma}}\right)$$

Substitute the expression for $V_n^{\hat{\gamma}}$ from the previous step:

$$-V_n^{\hat{\gamma}} = -\mu\Lambda X' + \frac{1}{2}X\Lambda X' - \frac{1}{2}X\hat{Q}^{-1}X'$$

Hence

$$C - V_n^{\hat{\gamma}} = \frac{1}{2}(\log|\Sigma| - \log|\hat{Q}|) + \frac{1}{2}\mu\Lambda\mu' - \mu\Lambda X' + \frac{1}{2}X\Lambda X' - \frac{1}{2}X\hat{Q}^{-1}X'$$

Under $\mathbb{P}$, the probability density function (pdf) of $X \sim \mathcal{N}(\mu, \Sigma)$ is

$$f_{\mathbb{P}}(x) = \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)\Lambda(x-\mu)'\right), \quad \text{with } \Lambda = \Sigma^{-1}$$

Similarly, the pdf of $X \sim \mathcal{N}(0, \hat{Q})$ is

$$f_{\hat{\mathbb{Q}}}(x) = \frac{1}{\sqrt{(2\pi)^n|\hat{Q}|}} \exp\left(-\frac{1}{2}x\hat{Q}^{-1}x'\right)$$

We want to see

$$\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}}(x) \equiv \frac{f_{\hat{\mathbb{Q}}}(x)}{f_{\mathbb{P}}(x)} = \frac{\exp\left(-\frac{1}{2}x\hat{Q}^{-1}x'\right)/\sqrt{(2\pi)^n|\hat{Q}|}}{\exp\left(-\frac{1}{2}(x-\mu)\Lambda(x-\mu)'\right)/\sqrt{(2\pi)^n|\Sigma|}}$$

Hence the ratio is:

$$\frac{f_{\hat{\mathbb{Q}}}(x)}{f_{\mathbb{P}}(x)} = \left[\frac{1}{\sqrt{(2\pi)^n|\hat{Q}|}}\middle/\frac{1}{\sqrt{(2\pi)^n|\Sigma|}}\right]\cdot\exp\left(-\frac{1}{2}x\hat{Q}^{-1}x' + \frac{1}{2}(x-\mu)\Lambda(x-\mu)'\right)$$

That front factor is simply $\sqrt{\frac{|\Sigma|}{|\hat{Q}|}}$. Next, expand the term $\frac{1}{2}(x-\mu)\Lambda(x-\mu)'$:

$$(x-\mu)\Lambda(x-\mu)' = x\Lambda x' - x\Lambda\mu' - \mu\Lambda x' + \mu\Lambda\mu'$$

Hence

$$-\frac{1}{2}x\hat{Q}^{-1}x' + \frac{1}{2}(x-\mu)\Lambda(x-\mu)' = -\frac{1}{2}x\hat{Q}^{-1}x' + \frac{1}{2}(x\Lambda x' - 2x\Lambda\mu' + \mu\Lambda\mu')$$

This is precisely the combination that shows up in $C - V_n^{\hat{\gamma}}$, which rearranges in a way that leaves you with:

$$\frac{1}{2}\mu\Lambda\mu' - \mu\Lambda x' + \frac{1}{2}x\Lambda x' - \frac{1}{2}x\hat{Q}^{-1}x'$$

Together with $\sqrt{\frac{|\Sigma|}{|\hat{Q}|}}$ out front, it matches the constant $\exp(C)$ factor as well. Therefore, by matching each term, one sees that

$$\exp(C - V_n^{\hat{\gamma}}(x)) = \frac{\exp\left(-\frac{1}{2}x\hat{Q}^{-1}x'\right)/\sqrt{(2\pi)^n|\hat{Q}|}}{\exp\left(-\frac{1}{2}(x-\mu)\Lambda(x-\mu)'\right)/\sqrt{(2\pi)^n|\Sigma|}}$$

From the above ratio, it follows that

$$\int_{\mathbb{R}^n}\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}}(x)f_{\mathbb{P}}(x)\,dx = \int_{\mathbb{R}^n}\frac{f_{\hat{\mathbb{Q}}}(x)}{f_{\mathbb{P}}(x)}f_{\mathbb{P}}(x)\,dx = \int_{\mathbb{R}^n}f_{\hat{\mathbb{Q}}}(x)\,dx = 1,$$

so indeed $\hat{\mathbb{Q}}$ is a probability measure. Moreover, $\hat{\mathbb{Q}}$ assigns to X precisely the density $f_{\hat{\mathbb{Q}}}$—i.e., the $\mathcal{N}(0,\hat{Q})$ distribution—by construction of that ratio. Hence, we have shown:

$$\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}}(x) = \frac{f_{\hat{\mathbb{Q}}}(x)}{f_{\mathbb{P}}(x)}, \quad \text{which implies} \quad (X;\hat{\mathbb{Q}}) \sim \mathcal{N}(0,\hat{Q})$$

Finally, we must verify that $\hat{\mathbb{Q}}$ belongs to the set $\mathcal{Q}_D$; that is, it must satisfy

$$\mathbb{E}_{\hat{\mathbb{Q}}}[S_t - S_s \mid \mathcal{G}_s^D] = 0 \quad \text{for all } t \geq s$$

Under $\hat{\mathbb{Q}}$, the vector X is distributed as $\mathcal{N}(0,\hat{Q})$. The observation is that the matrix $\hat{Q}$ has the property that its entries satisfy $\hat{Q}_{ij} = 0$ for $|i-j| > D$; this is inherent in the definition of the set $\mathcal{S}_D$. As a consequence, for each k, the increment X_k is independent of $X_1, \ldots, X_{(k-1-D)+}$ under $\hat{\mathbb{Q}}$. This conditional independence property implies, for any $s < t$, that

$$\mathbb{E}_{\hat{\mathbb{Q}}}[X_t \mid \mathcal{G}_s^D] = 0$$

Recalling that $S_t - S_s$ is the sum of the increments $X_{s+1} + \cdots + X_t$, linearity of expectation together with the independence implied by the bandedness of $\hat{Q}$ yields

$$\mathbb{E}_{\hat{\mathbb{Q}}}[S_t - S_s \mid \mathcal{G}_s^D] = 0$$

Thus, we conclude that $\hat{\mathbb{Q}} \in \mathcal{Q}_D$. By completing the above sub-steps, we have shown that the candidate strategy $\hat{\gamma}$ is admissible and, when paired with the measure $\hat{\mathbb{Q}}$ constructed via

$$\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}} = \exp(C - V_n^{\hat{\gamma}}),$$

satisfies the equality

$$V_n^{\hat{\gamma}} + \log\left(\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}}\right) = C$$

2.4 Summary of Chapter 2

This chapter has laid the theoretical foundation for optimal investment under delayed information. We have established that for an investor with exponential utility facing multivariate normal price increments and a fixed information delay D, there exists a unique optimal trading strategy. This strategy is explicitly given in terms of the mean price increments μ, the precision matrix $\Lambda = \Sigma^{-1}$, and a matrix $\hat{\Gamma}$ derived from a unique decomposition $\Lambda = \hat{Q}^{-1} + \hat{\Gamma}$. The matrix $\hat{Q}$ is D-banded, reflecting the information structure, while $\hat{\Gamma}$ captures adjustments due to the delay. A practical challenge in applying this theory is the determination of the precision matrix Λ and its decomposition for specific price process models. The next chapter explores this challenge.

Chapter 3

Precision Matrix for an MA(1)-like Price Process

The general theory developed in Chapter 2 highlights the central role of the precision matrix $\Lambda = \Sigma^{-1}$ of price increments. Obtaining this matrix can be difficult. This chapter delves into a specific, non-trivial example of a price process to illustrate the calculation of Σ and, more importantly, its inverse Λ.

3.1 The MA(1)-like Model for Prices

We consider a price process S_i constructed from underlying i.i.d. Gaussian random variables z_k.

In what follows we assume that

$$S_i = z_{i-1} + z_i, \quad i = 1, \dots, n,$$

where

$$z_0, z_1, \dots, z_n \sim \text{i.i.d. } \mathcal{N}(0, 1)$$

(In our setting the "initial" price S_0 is a constant; for convenience we may take $S_0 = 0$ because only increments matter.) We then define

$$X_i := S_i - S_{i-1}, \quad i = 1, \dots, n$$

A short calculation shows that

- For $i = 1$ we have
$$X_1 = S_1 - S_0 = z_0 + z_1,$$
- For $i \geq 2$ we have
$$X_i = S_i - S_{i-1} = (z_{i-1} + z_i) - (z_{i-2} + z_{i-1}) = z_i - z_{i-2}$$

3.2 Covariance Structure of Increments

The next step is to determine the covariance matrix Σ for the vector of price increments $\mathbf{X} = (X_1, \dots, X_n)'$.

Define the $(n+1)$-vector

$$\mathbf{z} = (z_0, z_1, \dots, z_n)',$$

and define the n-vector

$$\mathbf{X} = (X_1, X_2, \dots, X_n)'$$

Then one may write

$$\mathbf{X} = A\mathbf{z},$$

where the $n \times (n+1)$ matrix A is given by

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & 0 \\ -1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & 0 & 1 \end{pmatrix}.$$

For example, the second row is $(-1, 0, 1, 0, \dots, 0)$ since

$$X_2 = z_2 - z_0 = -1 \cdot z_0 + 0 \cdot z_1 + 1 \cdot z_2 + 0 \cdot z_3 + \dots)$$

Since the z_i's are independent with variance 1, the covariance matrix of $\mathbf{z}$ is the $(n+1) \times (n+1)$ identity, and hence the covariance matrix Σ of the increments $\mathbf{X}$ is

$$\Sigma := \operatorname{Cov}(\mathbf{X}) = AA'$$

We now compute the entries of Σ.

The diagonal entries

- For $i = 1$:

$$X_1 = z_0 + z_1 \quad \Longrightarrow \quad \operatorname{Var}(X_1) = \operatorname{Var}(z_0) + \operatorname{Var}(z_1) = 1 + 1 = 2$$

- For $i \geq 2$:

$$X_i = z_i - z_{i-2} \quad \Longrightarrow \quad \operatorname{Var}(X_i) = \operatorname{Var}(z_i) + \operatorname{Var}(z_{i-2}) = 1 + 1 = 2$$

Thus, for every $i = 1, \dots, n$,

$$\Sigma_{ii} = 2$$

The off-diagonal entries Let i and j be two indices from $\{1, \dots, n\}$.

1. When one of the indices is 1

   - For $i = 1, j \geq 2$, write

$$X_1 = z_0 + z_1, \quad X_j = \begin{cases} z_2 - z_0, & j = 2, \\ z_3 - z_1, & j = 3, \\ z_j - z_{j-2}, & j \geq 4 \end{cases}$$

   Then using independence we have:

– For $j = 2$: the only common term is z_0 (with coefficient $+1$ in X_1 and -1 in X_2); hence

$$\mathrm{Cov}(X_1, X_2) = 1 \cdot (-1) = -1.$$

– For $j = 3$: the only common term is z_1 (with coefficient $+1$ in X_1 and -1 in X_3); hence

$$\mathrm{Cov}(X_1, X_3) = -1$$

– For $j \geq 4$: there is no index k such that z_k appears in both X_1 and X_j; hence

$$\mathrm{Cov}(X_1, X_j) = 0$$

2. When $i, j \geq 2$.

- For $i, j \geq 2$ we have

$$X_i = z_i - z_{i-2} \quad \text{and} \quad X_j = z_j - z_{j-2}$$

Therefore,

$$\mathrm{Cov}(X_i, X_j) = \mathbb{E}[(z_i - z_{i-2})(z_j - z_{j-2})] = \delta_{ij} - \delta_{i,j-2} - \delta_{i-2,j} + \delta_{i-2,j-2},$$

where δ_{ab} is 1 if $a = b$ and 0 otherwise. In particular:

– If $i = j$ then

$$\mathrm{Cov}(X_i, X_i) = 1 + 1 = 2$$

– If $|i - j| = 2$ (with $i, j \geq 2$) then exactly one of the two "cross-terms" equals 1 and the others vanish, so

$$\mathrm{Cov}(X_i, X_j) = -1$$

– If $|i - j| \neq 0, 2$ then there is no overlap and

$$\mathrm{Cov}(X_i, X_j) = 0$$

Thus the $n \times n$ covariance matrix $\Sigma = (\Sigma_{ij})$ is given explicitly by

$$\Sigma_{ij} = \begin{cases} 2, & i = j, \\ -1, & \text{if } \{i, j\} = \{1, 2\} \text{ or } \{1, 3\}, \text{ or if } i, j \geq 2 \text{ and } |i - j| = 2, \\ 0, & \text{otherwise} \end{cases}$$

3.3 The Inverse Covariance (Precision) Matrix ${S^{(n)}}^{-1}$

For each integer $n \geq 3$ define the symmetric matrix $S^{(n)} \in \mathbb{R}^{n \times n}$

$$S^{(n)}_{ij} = \begin{cases} 2, & i = j, \\ -1, & \{i, j\} \in \{\{1, 2\}, \{1, 3\}\} \text{ or } (i, j \geq 2 \text{ and } |i - j| = 2), \\ 0, & \text{otherwise.} \end{cases}$$

Graph-theoretically $S^{(n)}$ is (identity + Laplacian) of the tree that consists of a **root vertex 1** to which two disjoint paths are attached:

- the even chain $\quad 2 - 4 - 6 - \cdots - (2n_e)$ of length $n_e = \lfloor n/2 \rfloor$;

- the odd chain $\quad 3 - 5 - 7 - \cdots - (2n_o + 1)$ of length $n_o = \lfloor (n-1)/2 \rfloor$.

Because $S^{(n)}$ is positive definite, it has a unique inverse $\Sigma^{(n)} = \left(S^{(n)}\right)^{-1}$.

The goal is a single explicit formula for every entry $\Sigma_{ij}^{(n)}$ and a formal proof that this formula is correct.

Throughout we set

$$L_e := n_e + 1, \qquad L_o := n_o + 1, \qquad \lambda := \frac{L_e L_o}{n+1}.$$

3.3.1 For the tridiagonal "path" matrix

$$F_m := \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}_{m \times m}$$

the inverse has a classical closed form:

$$(F_m^{-1})_{ij} = \frac{\min(i,j)\,(m+1-\max(i,j))}{m+1}, \qquad (1 \le i, j \le m). \tag{2.1}$$

Proof

Base $m = 1$. $F_1 = [2] \Rightarrow F_1^{-1} = [1/2]$ which matches (2.1).

Induction step. Write F_{m+1} in block form

$$F_{m+1} = \begin{bmatrix} 2 & e_1^\top \\ e_1 & F_m \end{bmatrix}, \quad e_1 = \begin{bmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

With the Schur complement $c = 2 - e_1^\top F_m^{-1} e_1 = 2 - \frac{m}{m+1} = \frac{m+2}{m+1}$ the block-inverse formula gives

$$F_{m+1}^{-1} = \begin{bmatrix} c^{-1} & -c^{-1} e_1^\top F_m^{-1} \\ -F_m^{-1} e_1 c^{-1} & F_m^{-1} + c^{-1} F_m^{-1} e_1 e_1^\top F_m^{-1} \end{bmatrix}$$

We prove that the closed form

$$(F_m^{-1})_{ij} = \frac{\min(i,j)\,(m+1-\max(i,j))}{m+1} \qquad (1 \le i, j \le m) \tag{2.1}$$

indeed gives the inverse of

$$F_m = \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}_{m \times m}$$

The base case $m = 1$ is immediate: $F_1 = [2]$ so $F_1^{-1} = [1/2]$.

Induction step $m \longrightarrow m+1$ Write

$$F_{m+1} = \begin{bmatrix} a & u^\top \\ u & B \end{bmatrix}, \quad a = 2, \quad u = \begin{bmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{m\times 1}, \quad B = F_m$$

Using the Schur complement, the block inverse is

$$F_{m+1}^{-1} = \begin{bmatrix} c^{-1} & -c^{-1}u^\top B^{-1} \\ -c^{-1}B^{-1}u & B^{-1} + c^{-1}B^{-1}uu^\top B^{-1} \end{bmatrix}, \qquad c = a - u^\top B^{-1} u \quad (\star)$$

Because u has only its first entry non-zero,

$$u^\top B^{-1} u = (B^{-1})_{11} = \frac{m}{m+1} \quad \text{(by 2.1)}, \qquad c = 2 - \frac{m}{m+1} = \frac{m+2}{m+1},$$

so

$$c^{-1} = \frac{m+1}{m+2} = \frac{\min(1,1)\,[m+2-\max(1,1)]}{m+2}. \tag{3.1}$$

Thus the $(1,1)$-entry matches (2.1).

For $j = 2, \dots, m+1$ let $j' := j-1 \in \{1, \dots, m\}$.
Using $(\star)$:

$$(F_{m+1}^{-1})_{1j} = -c^{-1}u^\top B^{-1} e_{j'} = c^{-1}(B^{-1})_{1j'} = \frac{m+1}{m+2} \cdot \frac{m+1-j'}{m+1} = \frac{m+2-j}{m+2}. \tag{3.2}$$

But $\min(1,j) = 1$ and $\max(1,j) = j$, so (3.2) coincides with (2.1).

Symmetry gives the entire first column.

Write $i' := i-1$, $j' := j-1 \in \{1, \dots, m\}$. From $(\star)$:

$$(F_{m+1}^{-1})_{ij} = (B^{-1})_{i'j'} + c^{-1}(B^{-1}u)_{i'}(u^\top B^{-1})_{j'}$$

$$= \frac{\min(i',j')(m+1-\max(i',j'))}{m+1} + \frac{m+1}{m+2} \cdot \frac{m+1-i'}{m+1} \cdot \frac{m+1-j'}{m+1} \tag{3.3}$$

Case A: $i' \le j'$ Then $\min(i',j') = i'$, $\max(i',j') = j'$. Using $m+1-j' = m+2-j$ and $i'+1 = i$,

$$\text{numerator of (3.3)} = i'(m+1-j') + \frac{m+1-i'}{m+2}(m+1-j')$$

$$= (m+1-j')\,[(m+2)i' + (m+1-i')] = (m+1-j')(m+1)(i'+1)$$

Dividing by $(m+1)(m+2)$ yields

$$(F_{m+1}^{-1})_{ij} = \frac{(i'+1)(m+1-j')}{m+2} = \frac{i(m+2-j)}{m+2}, \tag{3.4}$$

which is exactly (2.1) for $i \le j$.

Case B: $j' < i'$ The calculation is identical with i', j' swapped and produces

$$(F_{m+1}^{-1})_{ij} = \frac{j(m+2-i)}{m+2},$$

matching (2.1) for $j < i$. □

3.3.2 Block structure of $S^{(n)}$

Re-order the basis to $1 \mid 2, 4, 6, \cdots \mid 3, 5, 7, \ldots$.

Then

$$S^{(n)} = \begin{bmatrix} a & u^\top \\ u & B \end{bmatrix}, \quad a = 2, \quad u = \begin{bmatrix} -1 \\ 0 \\ \vdots \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad B = \operatorname{diag}(F_{n_e}, F_{n_o}). \tag{3.1}$$

Example for $n = 7$ In the natural numerical order $1, 2, 3, 4, 5, 6, 7$ the matrix $S^{(7)}$ looks "pentadiagonal":

$$S^{(7)} = \begin{pmatrix} 2 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 2 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 2 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 2 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 2 \end{pmatrix}$$

Re-label the vertices in the order

$$\pi = (1,\ 2, 4, 6,\ 3, 5, 7),$$

i.e. first the root, then the even chain, then the odd chain. Let P be the corresponding permutation matrix (the matrix whose columns are the standard basis vectors $e_{\pi(1)}, \ldots, e_{\pi(7)}$). Then

$$P\, S^{(7)}\, P^\top = \left[\begin{array}{c|ccc|ccc} 2 & -1 & 0 & 0 & -1 & 0 & 0 \\ \hline -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & 0 & 0 & 0 \\ \hline -1 & 0 & 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 \end{array}\right] = \begin{pmatrix} 2 & u^\top \\ u & \operatorname{diag}(F_3, F_3) \end{pmatrix}$$

Here $u = (-1, 0, 0, -1, 0, 0)^\top$ picks out the first vertex of each path, and each diagonal block F_3 is the familiar tridiagonal $2/-1$ "path" matrix defined in (2.1).

Block–inverse formula (Schur complement)

Let
$$M \;=\; \begin{bmatrix} a & u^\top \\ u & B \end{bmatrix}, \qquad a \in \mathbb{R},\; u \in \mathbb{R}^m,\; B \in \mathbb{R}^{m\times m},\; B \text{ invertible}$$

Because the upper–left block is the scalar a, the Schur complement is just the number

$$c \;=\; a - u^\top B^{-1} u.$$

Provided $c \neq 0$ (true in our application) one has

$$M^{-1} \;=\; \begin{bmatrix} c^{-1} & -c^{-1} u^\top B^{-1} \\ -c^{-1} B^{-1} u & B^{-1} + c^{-1} B^{-1} u\, u^\top B^{-1} \end{bmatrix}, \tag{*}$$

Verification: Multiply $(*)$ by M and check that all four block products give the identity; the key cancellations are $c^{-1}(a - u^\top B^{-1} u) = 1$ in the first row and $-c^{-1} B^{-1} u\, u^\top + B^{-1} u\, u^\top = 0$ in the lower block.

In our setting $a = 2$. Writing
$$\lambda \;:=\; \frac{1}{c} \;=\; \frac{1}{a - u^\top B^{-1} u}$$
and replacing c^{-1} by λ in $(*)$ gives the formula.

The inverse of such a block matrix is
$$\begin{bmatrix} \lambda & -\lambda u^\top B^{-1} \\ -\lambda B^{-1} u & B^{-1} + \lambda B^{-1} u u^\top B^{-1} \end{bmatrix}, \quad \lambda = \frac{1}{a - u^\top B^{-1} u} \tag{3.2}$$

Because only the *first* entry of each F_m^{-1} appears in $u^\top B^{-1} u$, formula (2.1) gives
$$u^\top B^{-1} u \;=\; \frac{n_e}{L_e} + \frac{n_o}{L_o} \;=\; 2 - \left(\tfrac{1}{L_e} + \tfrac{1}{L_o}\right)$$

Hence
$$2 - u^\top B^{-1} u \;=\; \frac{1}{L_e} + \frac{1}{L_o} \;=\; \frac{n+1}{L_e L_o}, \qquad \lambda \;=\; \frac{1}{2 - u^\top B^{-1} u} \;=\; \frac{L_e L_o}{n+1}$$

Finally we permute the inverse (3.2) back to the natural order $1, 2, 3, \ldots$ and express its entries in closed form. Let

$$\widehat{S} \;:=\; P\, S^{(n)}\, P^\top \;=\; \begin{bmatrix} 2 & u^\top \\ u & B \end{bmatrix}, \qquad B \;=\; \operatorname{diag}\Big(F_{n_e}, F_{n_o}\Big),$$

so that $\widehat{S}^{-1}$ is the block matrix produced by the Schur-complement calculation:

$$\widehat{S}^{-1} \;=\; \begin{bmatrix} \lambda & -\lambda u^\top B^{-1} \\ -\lambda B^{-1} u & B^{-1} + \lambda B^{-1} u\, u^\top B^{-1} \end{bmatrix}, \qquad \lambda \;=\; \frac{1}{2 - u^\top B^{-1} u}$$

Because a permutation matrix is orthogonal we have $P^\top = P^{-1}$; hence

$$\boxed{\Big(S^{(n)}\Big)^{-1} \;=\; P^\top\, \widehat{S}^{-1}\, P}$$

Example for $n = 7$ inversion Back to our example for $n = 7$

$$\widehat{S}^{-1} = \begin{pmatrix} 2 & \frac{3}{2} & 0 & 0 & \frac{3}{2} & 0 & 0 \\ \frac{3}{2} & \frac{15}{8} & \frac{9}{8} & 0 & 0 & 0 & 0 \\ 0 & \frac{9}{8} & \frac{15}{8} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{15}{8} & \frac{9}{8} & 0 & 0 \\ \frac{3}{2} & 0 & 0 & \frac{9}{8} & \frac{15}{8} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{7}{8} & \frac{1}{8} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{8} & \frac{7}{8} \end{pmatrix}.$$

Back to the natural order. Finally,

$$\left(S^{(7)}\right)^{-1} = P^{\top} \widehat{S}^{-1} P = \begin{pmatrix} 2 & \frac{3}{2} & \frac{3}{2} & 1 & 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{3}{2} & \frac{15}{8} & \frac{9}{8} & \frac{5}{4} & \frac{3}{4} & \frac{5}{8} & \frac{3}{8} \\ \frac{3}{2} & \frac{9}{8} & \frac{15}{8} & \frac{3}{4} & \frac{5}{4} & \frac{3}{8} & \frac{5}{8} \\ 1 & \frac{5}{4} & \frac{3}{4} & \frac{3}{2} & \frac{1}{2} & \frac{3}{4} & \frac{1}{4} \\ 1 & \frac{3}{4} & \frac{5}{4} & \frac{1}{2} & \frac{3}{2} & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{2} & \frac{5}{8} & \frac{3}{8} & \frac{3}{4} & \frac{1}{4} & \frac{7}{8} & \frac{1}{8} \\ \frac{1}{2} & \frac{3}{8} & \frac{5}{8} & \frac{1}{4} & \frac{3}{4} & \frac{1}{8} & \frac{7}{8} \end{pmatrix}$$

Statement of the problem

3.4 Explicit Formula for the Precision Matrix Entries

Leaf–distance function. For every internal vertex write

$$d(i) := \begin{cases} L_e - k, & i = 2k, \\ L_o - \ell, & i = 2\ell + 1 \ (\ell \geq 1), \end{cases}$$

so that $d(i)$ counts the number of edges from vertex i to the leaf of its own chain.

With this notation the inverse of $S^{(n)}$ is

$$\boxed{\Sigma_{ij}^{(n)} = \frac{1}{n+1} \begin{cases} L_e L_o, & i = j = 1, \\ L_o \, d(j), & i = 1, \ j \text{ even}, \\ L_e \, d(j), & i = 1, \ j \text{ odd} > 1, \\ d(i) \, d(j), & i \text{ even}, \ j \text{ odd}, \\ \dfrac{n+1}{L_e} \min(k, \ell)\Big(L_e - \max(k, \ell)\Big) + \dfrac{L_o}{L_e} d(i) \, d(j), & i = 2k, \ j = 2\ell, \\ \dfrac{n+1}{L_o} \min(k, \ell)\Big(L_o - \max(k, \ell)\Big) + \dfrac{L_e}{L_o} d(i) \, d(j), & i = 2k+1, \ j = 2\ell + 1. \end{cases}}$$

We already proved (2.1) for the path inverse by induction. To elevate that result to $S^{(n)}$ itself, induct from n to $n + 2$:

- appending two new leaf vertices $n+1, n+2$ (one on each chain) enlarges B by one row and column in each diagonal block, whose inverse is given by the induction hypothesis (2.1). After the extension we have $L_e \mapsto L_e + 1$ and $L_o \mapsto L_o + 1$; consequently the factor in front of the matrix changes from $\frac{1}{n+1}$ to $\frac{1}{n+3}$.;
- the update of u merely appends two zeros, so the scalar λ in (3.2) changes from $\lambda_n = \frac{L_e L_o}{n+1}$ to $\lambda_{n+2} = \frac{(L_e+1)(L_o+1)}{n+3}$;
- direct substitution shows that (4.1) with the enlarged parameters reproduces the new block-inverse (3.2).

Because the base case $n = 3$ coincides with $\Sigma^{(3)} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{3}{4} \end{bmatrix}$, formula (4.1) holds for **all** $n \geq 3$.

■

3.5 Verification and Implications

Checking if this is indeed the inverse leveraging Python and the SymPy library

```
import sympy as sp

def _parameters(n: int):
    """Return (n_e, n_o, L_e, L_o) for a given n  3."""
    if n < 3:
        raise ValueError("n must be at least 3.")
    n_e = n // 2                    # n/2
    n_o = (n - 1) // 2              # (n1)/2
    L_e = n_e + 1
    L_o = n_o + 1
    return n_e, n_o, L_e, L_o


def S_matrix(n: int) -> sp.Matrix:
    """Precision matrix S^{(n)}:

        S_{ij}  =  2,     i = j
                = -1,    {i,j} = {1,2} or {1,3} or (i, j  2 and |i-j| = 2)
                =  0,    otherwise
    """
    n_e, n_o, L_e, L_o = _parameters(n)
    S = sp.zeros(n, n)

    for i in range(n):
        for j in range(n):
            if i == j:
                S[i, j] = 2
            elif (i == 0 and j in (1, 2)) or (j == 0 and i in (1, 2)):
                S[i, j] = -1
            elif i >= 1 and j >= 1 and abs(i - j) == 2:
                S[i, j] = -1
            else:
                S[i, j] = 0
    return S
```

```
def _leaf_distance(index: int, n: int, L_e: int, L_o: int):
    i = index
    if i == 1:
        return None  # not used for the root
    if i
        k = i // 2
        return L_e - k
    else:           # odd vertex 2+1,   1
         = (i - 1) // 2
        return L_o -


def sigma_entry(i: int, j: int, n: int) -> sp.Rational:
    n_e, n_o, L_e, L_o = _parameters(n)
    factor = sp.Rational(1, n + 1)

    # Symmetry
    if j < i:
        return sigma_entry(j, i, n)

    # Root row/column
    if i == 1 and j == 1:
        return factor * L_e * L_o

    if i == 1:
        d_j = _leaf_distance(j, n, L_e, L_o)
        if j
            return factor * L_o * d_j
        else:                  # j odd (>1)
            return factor * L_e * d_j

    # Mixed parity
    if (i
        d_i = _leaf_distance(i, n, L_e, L_o)
        d_j = _leaf_distance(j, n, L_e, L_o)
        return factor * d_i * d_j

    # Same parity
    if i
        k = i // 2
         = j // 2
        d_i = L_e - k
        d_j = L_e -
        term1 = sp.Rational(n + 1, L_e) * min(k, ) * (L_e - max(k, ))
        term2 = sp.Rational(L_o, L_e) * d_i * d_j
        return factor * (term1 + term2)

    else:            # both odd (3) : i = 2k+1, j = 2+1
        k = (i - 1) // 2
         = (j - 1) // 2
        d_i = L_o - k
        d_j = L_o -
        term1 = sp.Rational(n + 1, L_o) * min(k, ) * (L_o - max(k, ))
        term2 = sp.Rational(L_e, L_o) * d_i * d_j
        return factor * (term1 + term2)


def Sigma_matrix(n: int) -> sp.Matrix:
```

```
     = sp.zeros(n, n)
    for i in range(1, n + 1):
        for j in range(1, n + 1):
            [i - 1, j - 1] = sigma_entry(i, j, n)
    return


# Demo / verification
if __name__ == "__main__":
    n_demo = 7
    S = S_matrix(n_demo)
    _explicit = Sigma_matrix(n_demo)

    # Check: S * \Sigma == I
    is_identity = (S * _explicit) == sp.eye(n_demo)

    print("S^(n) for n =", n_demo)
    sp.pprint(S)
    print("\n^(n) (explicit) for n =", n_demo)
    sp.pprint(_explicit)
    print("\nVerification S *  = I ? ->", is_identity)
```

$$
\underbrace{\begin{pmatrix} 2 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 2 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 2 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 2 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 2 \end{pmatrix}}_{S}
\underbrace{\begin{pmatrix} 2 & \frac{3}{2} & \frac{3}{2} & 1 & 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{3}{2} & \frac{15}{8} & \frac{9}{8} & \frac{5}{4} & \frac{3}{4} & \frac{5}{8} & \frac{3}{8} \\ \frac{3}{2} & \frac{9}{8} & \frac{15}{8} & \frac{3}{4} & \frac{5}{4} & \frac{3}{8} & \frac{5}{8} \\ 1 & \frac{5}{4} & \frac{3}{4} & \frac{3}{2} & \frac{1}{2} & \frac{3}{4} & \frac{1}{4} \\ 1 & \frac{3}{4} & \frac{5}{4} & \frac{1}{2} & \frac{3}{2} & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{2} & \frac{5}{8} & \frac{3}{8} & \frac{3}{4} & \frac{1}{4} & \frac{7}{8} & \frac{1}{8} \\ \frac{1}{2} & \frac{3}{8} & \frac{5}{8} & \frac{1}{4} & \frac{3}{4} & \frac{1}{8} & \frac{7}{8} \end{pmatrix}}_{S^{-1}}
\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}}_{S\,S^{-1}}
$$

Numerical illustration of the optimal–weight kernel

The closed forms of Section 3 yield an explicit $\hat{\gamma}$–kernel, but the algebra is lengthy and not enlightening. Instead we evaluate the formula symbolically in `Python/SymPy` and visualise the resulting coefficients.

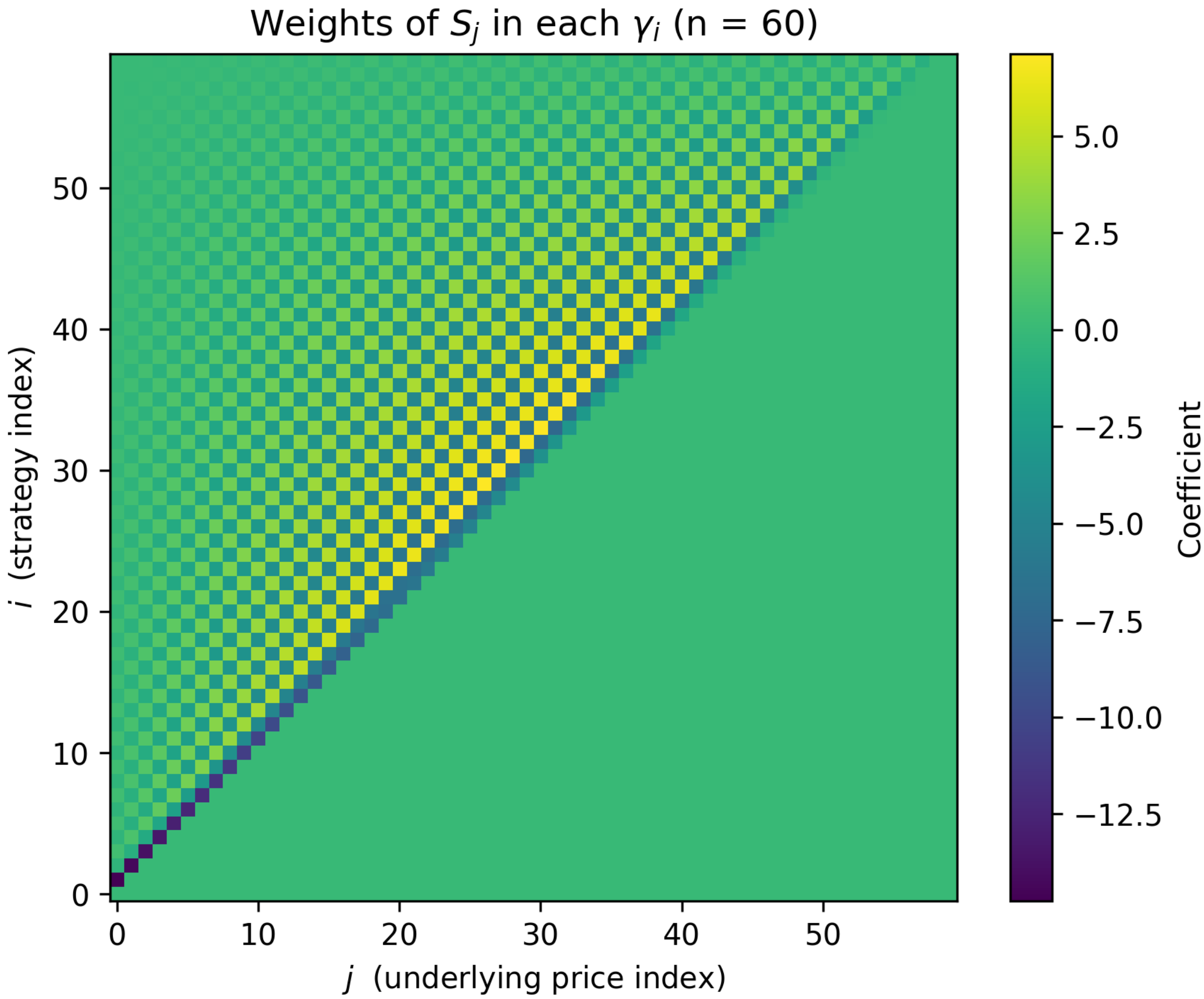

Figure 3.1: Heat-map of the weights that every past price level S_j receives inside each trading decision γ_i when the delay is $D = 0$ and the horizon is $n = 60$. Row i ($i = 1, \dots, n$) corresponds to the strategy coefficient vector of γ_i; column j ($j < i$) corresponds to the underlying price S_j. The colour encodes the net coefficient $C_{ij} = -\hat{\Gamma}_{ij} + \mathbf{1}_{\{j+1<i\}}\hat{\Gamma}_{i,j+1}$, so that a positive (yellow) tile means an upward move in S_j increases γ_i, whereas a negative (purple/blue) tile decreases it. The plot is strictly upper-triangular because each γ_i only depends on prices observed up to time $i-1$. The alternating sign pattern reflects the telescoping form of the increments $S_j - S_{j-1}$ and the banded structure of the inverse covariance matrix $\Lambda = S^{-1}$.

Chapter 4

Optimal Investment in an Autoregressive (AR(1)) Market

Road-map and connection to previous chapters

The present chapter serves a different purpose than the previous ones:

1. It treats the standard AR(1) model

$$X_t \;=\; (1+\beta)\, X_{t-1} + \sigma\, \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0,1),$$

for which fully-informed optimal policies are already documented in (Deák and Rásonyi, 2015).

2. By re-computing the precision matrix for the increments $\Delta X_t := X_t - X_{t-1}$ and feeding it into the general formula of Chapter 2, we verify that our delayed-information theory collapses to the already proven solution when the delay is $D = 0$. This completes the logical circle announced in the Introduction.

3. For simplicity, I will be referring to Deák and Rásonyi (2015) as the second paper and Dolinsky and Zuk (2023) as the first paper.

4.1 Model set-up

The second paper considers a single risky asset $(X_t)_{t\geq 0}$ evolving as:

$$X_t = (1+\beta)X_{t-1} + \sigma\epsilon_t, \quad t \geq 1,$$

with $X_0 = z$ given, $\beta \in \mathbb{R}$, $\sigma > 0$, and $\epsilon_t \sim i.i.d.\ \mathcal{N}(0,1)$.

Define the increments:

$$\Delta X_t = X_t - X_{t-1}$$

From the model:

$$\Delta X_t = \beta X_{t-1} + \sigma\epsilon_t$$

4.2 Covariance and precision matrix of the increments

We know:

$$X_{t-1} = z(1+\beta)^{t-1} + \sigma \sum_{k=1}^{t-1} (1+\beta)^{t-1-k}\epsilon_k$$

Thus:

$$\Delta X_t = \beta z(1+\beta)^{t-1} + \sigma\epsilon_t + \beta\sigma\sum_{k=1}^{t-1}(1+\beta)^{t-1-k}\epsilon_k$$

To isolate the random part, define the vector $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_T)^\top$. We can write:

$$\Delta X_t = \underbrace{\beta z(1+\beta)^{t-1}}_{\mu_t} + \sum_{k=1}^{t} L_{t,k}\epsilon_k,$$

where the coefficients $L_{t,k}$ are determined as follows:

- For $k = t$: The coefficient of ϵ_t is σ.
- For $1 \le k < t$: The coefficient of ϵ_k is $\beta\sigma(1+\beta)^{t-1-k}$.
- For $k > t$: No dependence, so $L_{t,k} = 0$.

This gives the random part of ΔX_t as:

$$\Delta X_t - \mu_t = \sigma\epsilon_t + \beta\sigma\sum_{k=1}^{t-1}(1+\beta)^{t-1-k}\epsilon_k$$

In matrix form, let $\Delta\boldsymbol{X} = (\Delta X_1, \ldots, \Delta X_T)^\top$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_T)^\top$. Define the lower-triangular matrix $L = (L_{t,k})_{1\le k\le T}$ by:

$$L_{t,t} = \sigma, \quad L_{t,k} = \begin{cases} \beta\sigma(1+\beta)^{t-1-k}, & k < t, \\ 0, & k > t. \end{cases}$$

Hence:

$$\Delta\boldsymbol{X} = \boldsymbol{\mu} + L\boldsymbol{\epsilon}$$

Since $\boldsymbol{\epsilon} \sim \mathcal{N}(0, I_T)$, the covariance matrix of $\Delta\boldsymbol{X}$ is:

$$\Sigma = \mathrm{Var}(\Delta\boldsymbol{X}) = \mathrm{Var}(L\boldsymbol{\epsilon}) = LL^\top$$

Explicit Covariance Matrix

Thus Σ is explicitly:

$$\Sigma_{t,s} = \sum_{k=1}^{\min(t,s)} L_{t,k}L_{s,k}$$

Substituting $L_{t,k}$:

- If $t = s$:

$$\Sigma_{t,t} = L_{t,t}^2 + \sum_{k=1}^{t-1}[\beta\sigma(1+\beta)^{t-1-k}]^2 = \sigma^2\left[1 + \beta^2\sum_{i=0}^{t-2}(1+\beta)^{2i}\right].$$

- If $t > s$:

$$\Sigma_{t,s} = L_{t,s}L_{s,s} + \sum_{k=1}^{s-1} L_{t,k}L_{s,k}.$$

Concrete Example for $T = 3$

Step 1: Construct L

The lower triangular matrix L is constructed as follows:

$$L = \begin{pmatrix} L_{1,1} & 0 & 0 \\ L_{2,1} & L_{2,2} & 0 \\ L_{3,1} & L_{3,2} & L_{3,3} \end{pmatrix} = \begin{pmatrix} \sigma & 0 & 0 \\ \beta\sigma & \sigma & 0 \\ \beta\sigma(1+\beta) & \beta\sigma & \sigma \end{pmatrix}$$

Step 2: Compute $\Sigma = LL^\top$

The entries of Σ are computed as:

$$\begin{aligned}
&\Sigma_{1,1} = \sigma^2, \\
&\Sigma_{2,1} = \beta\sigma \cdot \sigma = \beta\sigma^2, \\
&\Sigma_{2,2} = (\beta\sigma)^2 + \sigma^2 = \sigma^2(\beta^2 + 1), \\
&\Sigma_{3,1} = [\beta\sigma(1+\beta)]\sigma = \beta(1+\beta)\sigma^2, \\
&\Sigma_{3,2} = [\beta\sigma(1+\beta)](\beta\sigma) + (\beta\sigma)\sigma = \beta^2(1+\beta)\sigma^2 + \beta\sigma^2 = \beta\sigma^2[\beta(1+\beta)+1], \\
&\Sigma_{3,3} = [\beta\sigma(1+\beta)]^2 + (\beta\sigma)^2 + \sigma^2 = \beta^2\sigma^2(1+\beta)^2 + \beta^2\sigma^2 + \sigma^2 = \sigma^2[1 + \beta^2 + \beta^2(1+\beta)^2].
\end{aligned}$$

Thus, for $T = 3$, Σ is:

$$\Sigma = \begin{pmatrix} \sigma^2 & \beta\sigma^2 & \beta(1+\beta)\sigma^2 \\ \beta\sigma^2 & \sigma^2(\beta^2+1) & \beta\sigma^2[\beta(1+\beta)+1] \\ \beta(1+\beta)\sigma^2 & \beta\sigma^2[\beta(1+\beta)+1] & \sigma^2[1+\beta^2+\beta^2(1+\beta)^2] \end{pmatrix}$$

The mean vector $\boldsymbol{\mu}$ is:

$$\boldsymbol{\mu} = \begin{pmatrix} \beta z \\ \beta z(1+\beta) \\ \beta z(1+\beta)^2 \end{pmatrix}$$

Step 1: Inverting L

To find Σ^{-1}, we first invert L. Let $M = L^{-1}$. Since L is lower-triangular with σ on the diagonal, M will be lower-triangular with $M_{t,t} = 1/\sigma$. The off-diagonal elements are determined by the requirement $ML = I$. For example:

- For $T = 1$:
$$L = [\sigma], \quad M = [1/\sigma].$$
- For $T = 2$:
$$L = \begin{pmatrix} \sigma & 0 \\ \beta\sigma & \sigma \end{pmatrix}, \quad M = \begin{pmatrix} 1/\sigma & 0 \\ -\beta/\sigma & 1/\sigma \end{pmatrix}.$$
- For $T = 3$:
$$L = \begin{pmatrix} \sigma & 0 & 0 \\ \beta\sigma & \sigma & 0 \\ \beta\sigma(1+\beta) & \beta\sigma & \sigma \end{pmatrix}, \quad M = \begin{pmatrix} 1/\sigma & 0 & 0 \\ -\beta/\sigma & 1/\sigma & 0 \\ -\beta/\sigma & -\beta/\sigma & 1/\sigma \end{pmatrix}.$$

For general T, $M = L^{-1}$ is:

$$M_{t,t} = \frac{1}{\sigma}, \quad M_{t,k} = \begin{cases} -\beta/\sigma, & \text{if } k < t, \\ 0, & \text{if } k > t. \end{cases}$$

Step 2: Finding Σ^{-1}

We know $\Sigma = LL^\top$, so:

$$\Sigma^{-1} = (L^\top)^{-1} L^{-1} = M^\top M$$

The matrix M is:

$$M = \begin{pmatrix} 1/\sigma & 0 & 0 & \cdots & 0 \\ -\beta/\sigma & 1/\sigma & 0 & \cdots & 0 \\ -\beta/\sigma & -\beta/\sigma & 1/\sigma & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\beta/\sigma & -\beta/\sigma & -\beta/\sigma & \cdots & 1/\sigma \end{pmatrix}$$

Each column t of M has the following structure:

- Zeros in rows $1, \dots, t-1$.
- A $1/\sigma$ at row t.
- $-\beta/\sigma$ in rows $t+1, \dots, T$.

Step 3: The Entries of Σ^{-1}

Let $\Gamma = \Sigma^{-1} = M^\top M$. To find $\Gamma_{t,s}$ for $1 \le t \le s \le T$:

$$\Gamma_{t,s} = (M^\top M)_{t,s} = \sum_{k=1}^{T} M_{k,t} M_{k,s}$$

- **Diagonal entries ($t = s$):** For the t-th diagonal element $\Gamma_{t,t}$:

$$\Gamma_{t,t} = \frac{1}{\sigma^2} + (T - t)\frac{\beta^2}{\sigma^2} = \frac{1 + \beta^2(T - t)}{\sigma^2}.$$

- **Off-diagonal entries ($t \neq s$):** These can be computed similarly by summing over the non-zero overlaps of M.

Final Closed-Form for Σ^{-1}

We have derived that for all $1 \le t \le T$:

$$(\Sigma^{-1})_{t,t} = \frac{1 + \beta^2(T - t)}{\sigma^2}$$

For $1 \le t < s \le T$:

$$(\Sigma^{-1})_{t,s} = (\Sigma^{-1})_{s,t} = \frac{\beta^2(T - s) - \beta}{\sigma^2}$$

This gives us every entry of the precision matrix Σ^{-1} explicitly in terms of β, σ, and the time indices t, s.

4.3 Deriving the optimal strategy via the general theorem

We now take

$$\hat{\gamma}_i = \underbrace{\sum_{j=1}^{T} \Lambda_{i,j}\mu_j}_{\text{Term A}} - \underbrace{\sum_{j=1}^{i-1} \Lambda_{i,j}\Delta X_j}_{\text{Term B}}, \quad i = 1, \ldots, T,$$

and insert

- $\mu_j = \beta z(1+\beta)^{j-1}$,
- $\Delta X_j = \beta X_{j-1} + \sigma\epsilon_j$,
- $\Lambda_{i,j}$ from the explicit entries of Σ^{-1}.

We do this in two big parts: **Term A** then **Term B**, and sum them.

4.3.1 Term A: $\sum_{j=1}^{T} \Lambda_{i,j}\mu_j$

We write

$$\text{Term A} = \sum_{j=1}^{T} \Lambda_{i,j}\beta z(1+\beta)^{j-1}$$

Depending on whether $j \leq i$ or $j > i$, the expression for $\Lambda_{i,j}$ changes (diagonal vs. off-diagonal). Let us separate the sum:

$$\sum_{j=1}^{T} = \sum_{j=1}^{i-1} + \sum_{j=i}^{i} + \sum_{j=i+1}^{T}$$

Since the middle sum is just the single diagonal term $j = i$, we split:

$$\text{Term A} = \sum_{j=1}^{i-1} \Lambda_{i,j}\beta z(1+\beta)^{j-1} + \underbrace{\Lambda_{i,i}\beta z(1+\beta)^{i-1}}_{\text{diagonal part}} + \sum_{j=i+1}^{T} \Lambda_{i,j}\beta z(1+\beta)^{j-1}$$

Using the formula for $\Lambda_{i,j}$:

- **Diagonal ($i = j$):**

$$\Lambda_{i,i} = \frac{1+\beta^2(T-i)}{\sigma^2}.$$

- **Off-diagonal:**
  - For $j < i$, we use $\Lambda_{i,j} = \Lambda_{j,i}$:

$$\Lambda_{i,j} = \frac{\beta^2(T-i)-\beta}{\sigma^2}, \quad \text{for } 1 \leq j < i.$$

  - For $j > i$:

$$\Lambda_{i,j} = \frac{\beta^2(T-j)-\beta}{\sigma^2}.$$

Thus:

$$\text{Term A} = \sum_{j=1}^{i-1} \left[\frac{\beta^2(T-i) - \beta}{\sigma^2} \right] \beta z (1+\beta)^{j-1} + \left[\frac{1 + \beta^2(T-i)}{\sigma^2} \right] \beta z (1+\beta)^{i-1} +$$

$$+ \sum_{j=i+1}^{T} \left[\frac{\beta^2(T-j) - \beta}{\sigma^2} \right] \beta z (1+\beta)^{j-1}$$

4.3.2 Term B: $-\sum_{j=1}^{i-1} \Lambda_{i,j} \Delta X_j$

Recall that

$$\Delta X_j = \beta X_{j-1} + \sigma \epsilon_j$$

Thus,

$$\text{Term B} = -\sum_{j=1}^{i-1} \Lambda_{i,j} \left[\beta X_{j-1} + \sigma \epsilon_j \right]$$

Since $j \leq i - 1 < i$, we use the formula:

$$\Lambda_{i,j} = \frac{\beta^2(T-i) - \beta}{\sigma^2}, \quad \text{for all } j = 1, \ldots, i-1$$

Hence,

$$\text{Term B} = -\sum_{j=1}^{i-1} \left[\frac{\beta^2(T-i) - \beta}{\sigma^2} \right] \left[\beta X_{j-1} + \sigma \epsilon_j \right]$$

Factoring out $\frac{\beta^2(T-i)-\beta}{\sigma^2}$:

$$\text{Term B} = -\left[\frac{\beta^2(T-i) - \beta}{\sigma^2} \right] \sum_{j=1}^{i-1} \left[\beta X_{j-1} + \sigma \epsilon_j \right]$$

4.4 Combine Term A and Term B

Putting these together,

$$\hat{\gamma}_i = \text{Term A} + \text{Term B} = \left[(\star_1) + (\star_2) + (\star_3) \right] - \left[\frac{\beta^2(T-i) - \beta}{\sigma^2} \right] \sum_{j=1}^{i-1} \left[\beta X_{j-1} + \sigma \epsilon_j \right]$$

Recalling:

$$(\star_1) = \sum_{j=1}^{i-1} \left[\frac{\beta^2(T-i) - \beta}{\sigma^2} \right] \beta z (1+\beta)^{j-1},$$

$$(\star_2) = \left[\frac{1 + \beta^2(T-i)}{\sigma^2} \right] \beta z (1+\beta)^{i-1},$$

$$(\star_3) = \sum_{j=i+1}^{T} \left[\frac{\beta^2(T-j) - \beta}{\sigma^2} \right] \beta z (1+\beta)^{j-1}.$$

We see that $(\star_1)$ and Term B share the same factor $\frac{\beta^2(T-i)-\beta}{\sigma^2}$. Grouping terms:

$$(\star_1) - \left[\frac{\beta^2(T-i) - \beta}{\sigma^2} \right] \sum_{j=1}^{i-1} \beta X_{j-1} = \left[\frac{\beta^2(T-i) - \beta}{\sigma^2} \right] \left(\beta \sum_{j=1}^{i-1} z(1+\beta)^{j-1} - \beta \sum_{j=1}^{i-1} X_{j-1} \right)$$

Factoring:

$$= \left[\frac{\beta^2(T-i) - \beta}{\sigma^2} \right] \beta \sum_{j=1}^{i-1} \left[z(1+\beta)^{j-1} - X_{j-1} \right]$$

Handling the ϵ_j terms:

$$- \left[\frac{\beta^2(T-i) - \beta}{\sigma^2} \right] \sum_{j=1}^{i-1} \sigma \epsilon_j$$

Final expression:

$$\hat{\gamma}_i = \left[\frac{\beta^2(T-i) - \beta}{\sigma^2} \right] \beta \sum_{j=1}^{i-1} \left[z(1+\beta)^{j-1} - X_{j-1} \right] + (\star_2) + (\star_3) - \left[\frac{\beta^2(T-i) - \beta}{\sigma^2} \right] \sum_{j=1}^{i-1} \sigma \epsilon_j$$

4.5 Expanding All Pieces for the AR(1) Model

After expanding all pieces for the AR(1) model,

$$X_t = (1+\beta) X_{t-1} + \sigma \epsilon_t, \quad \Delta X_t = X_t - X_{t-1} = \beta X_{t-1} + \sigma \epsilon_t,$$

we ended up with:

$$\hat{\gamma}_i = \underbrace{\left[\frac{\beta^2(T-i) - \beta}{\sigma^2} \right] \beta \sum_{j=1}^{i-1} \left[z(1+\beta)^{j-1} - X_{j-1} \right]}_{\text{(A)}} + \underbrace{\left[\beta \frac{1 + \beta^2(T-i)}{\sigma^2} \right] z(1+\beta)^{i-1}}_{\text{(B)}}$$

$$+ \underbrace{\sum_{j=i+1}^{T} \left[\frac{\beta^2(T-j) - \beta}{\sigma^2} \right] \beta z (1+\beta)^{j-1}}_{\text{(C)}} - \underbrace{\left[\frac{\beta^2(T-i) - \beta}{\sigma^2} \right] \sum_{j=1}^{i-1} \sigma \epsilon_j}_{\text{(D)}}$$

4.5.1 Outline of the Plan

The expression for $\hat{\gamma}_i$ above has **four** parts: (A), (B), (C), (D). Notice:

1. **Parts (A) and (D)** each contain the common factor

$$\frac{\beta^2(T-i)-\beta}{\sigma^2}.$$

   Part (A) multiplies it by a sum of $[z(1+\beta)^{j-1} - X_{j-1}]$, and part (D) multiplies it by a sum of $\sigma\epsilon_j$.

2. **Part (B)** is a single "diagonal" term involving $z(1+\beta)^{i-1}$.

3. **Part (C)** is a sum from $j = i+1$ to T, also involving $z(1+\beta)^{j-1}$.

Crucially, X_{i-1} itself (for an AR(1)) can be written in closed form:

$$X_{i-1} = z(1+\beta)^{i-1} + \sigma \sum_{m=1}^{i-1} (1+\beta)^{i-1-m}\epsilon_m$$

4.6 Detailed Step-by-Step Telescopes

We want to handle terms (A), (B), (C), (D) in a systematic way.

4.6.1 Rewrite $z(1+\beta)^{j-1} - X_{j-1}$

Focus on the sum inside (A):

$$\sum_{j=1}^{i-1} \left[z(1+\beta)^{j-1} - X_{j-1}\right]$$

For each j,

$$X_{j-1} = z(1+\beta)^{j-1} + \sigma \sum_{m=1}^{j-1} (1+\beta)^{j-1-m}\epsilon_m$$

Thus,

$$z(1+\beta)^{j-1} - X_{j-1} = -\sigma \sum_{m=1}^{j-1} (1+\beta)^{j-1-m}\epsilon_m$$

Therefore, part (A) becomes:

$$(\mathrm{A}) = \left[\frac{\beta^2(T-i)-\beta}{\sigma^2}\right] \beta \sum_{j=1}^{i-1} \left[-\sigma \sum_{m=1}^{j-1} (1+\beta)^{j-1-m}\epsilon_m\right]$$

Factoring out the minus sign and σ:

$$(\text{A}) = -\left[\frac{\beta^2(T-i)-\beta}{\sigma^2}\right]\beta\sigma\sum_{j=1}^{i-1}\sum_{m=1}^{j-1}(1+\beta)^{j-1-m}\epsilon_m$$

Interchanging the summation order:

$$(\text{A}) = -\left[\frac{\beta^2(T-i)-\beta}{\sigma^2}\right]\beta\sigma\sum_{m=1}^{i-2}\sum_{j=m+1}^{i-1}(1+\beta)^{j-1-m}\epsilon_m$$

The inner sum is a finite geometric series:

$$\sum_{j=m+1}^{i-1}(1+\beta)^{j-1-m} = \frac{(1+\beta)^{i-1-m}-1}{\beta}$$

Thus,

$$(\text{A}) = -\left[\frac{\beta^2(T-i)-\beta}{\sigma^2}\right]\sigma\sum_{m=1}^{i-2}\left[(1+\beta)^{i-1-m}-1\right]\epsilon_m$$

$$= -\left[\frac{\beta^2(T-i)-\beta}{\sigma^2}\right]\sigma\sum_{m=1}^{i-2}(1+\beta)^{i-1-m}\epsilon_m + \left[\frac{\beta^2(T-i)-\beta}{\sigma^2}\right]\sigma\sum_{m=1}^{i-2}\epsilon_m$$

Finally, writing it more compactly:

$$(\text{A}) = -\frac{\beta^2(T-i)-\beta}{\sigma^2}\sigma\sum_{m=1}^{i-1}(1+\beta)^{i-1-m}\epsilon_m$$

4.6.2 Compare Part (A) with Part (D)

Part (D) was:

$$(\text{D}) = -\left[\frac{\beta^2(T-i)-\beta}{\sigma^2}\right]\sum_{j=1}^{i-1}\sigma\epsilon_j$$

So combining (A) + (D):

$$(\text{A}) + (\text{D}) = -\frac{\beta^2(T-i)-\beta}{\sigma^2}\sigma\sum_{m=1}^{i-1}(1+\beta)^{i-1-m}\epsilon_m + \sum_{j=1}^{i-1}\sigma\epsilon_j$$

Rewriting the sum as a single term:

$$(\text{A}) + (\text{D}) = -\frac{\beta^2(T-i)-\beta}{\sigma^2}\sigma\sum_{m=1}^{i-1}(1+\beta)^{i-1-m}\epsilon_m$$

$$\boxed{(\text{A}) + (\text{D}) = -\frac{\beta^2(T-i)-\beta}{\sigma^2}\sigma\sum_{m=1}^{i-1}(1+\beta)^{i-1-m}\epsilon_m.}$$

4.6.3 Rewrite That ϵ-Sum in Terms of X_{i-1}

Recall the AR(1) expression for X_{i-1}:

$$X_{i-1} = z(1+\beta)^{i-1} + \sigma \sum_{m=1}^{i-1} (1+\beta)^{i-1-m} \epsilon_m$$

Hence,

$$\sigma \sum_{m=1}^{i-1} (1+\beta)^{i-1-m} \epsilon_m = X_{i-1} - z(1+\beta)^{i-1}$$

Thus,

$$(\text{A}) + (\text{D}) = -\frac{\beta^2(T-i)-\beta}{\sigma^2} \left[X_{i-1} - z(1+\beta)^{i-1}\right]$$

Expanding,

$$(\text{A}) + (\text{D}) = -\frac{\beta^2(T-i)-\beta}{\sigma^2} X_{i-1} + \frac{\beta^2(T-i)-\beta}{\sigma^2} z(1+\beta)^{i-1}$$

So we rewrite:

$$\boxed{(\text{A}) + (\text{D}) = -\frac{\beta^2(T-i)-\beta}{\sigma^2} X_{i-1} + \frac{\beta^2(T-i)-\beta}{\sigma^2} z(1+\beta)^{i-1}.}$$

4.6.4 Add Parts (B) and (C)

Now recall:

$$(\text{B}) = \frac{\beta(1+\beta^2(T-i))}{\sigma^2} z(1+\beta)^{i-1}$$

$$(\text{C}) = \sum_{j=i+1}^{T} \left[\frac{\beta^2(T-j)-\beta}{\sigma^2}\right] \beta z(1+\beta)^{j-1}$$

So,

$$(\text{B}) = \frac{\beta(1+\beta^2(T-i))}{\sigma^2} z(1+\beta)^{i-1}, \quad (\text{C}) = \beta z \sum_{j=i+1}^{T} \frac{\beta^2(T-j)-\beta}{\sigma^2} (1+\beta)^{j-1}$$

4.6.5 Combine All Four Parts: (A) + (B) + (C) + (D)

From 3.3 we have

$$(\text{A}) + (\text{D}) = -\frac{\beta^2(T-i)-\beta}{\sigma^2} X_{i-1} + \frac{\beta^2(T-i)-\beta}{\sigma^2} z(1+\beta)^{i-1}$$

Hence,

$$\hat{\gamma}_i = \text{(A)} + \text{(B)} + \text{(C)} + \text{(D)}$$

So,

$$\hat{\gamma}_i = -\frac{\beta^2(T-i)-\beta}{\sigma^2}X_{i-1} + \underbrace{\frac{\beta^2(T-i)-\beta}{\sigma^2}z(1+\beta)^{i-1}}_{\text{call this (E1)}} + \underbrace{\frac{\beta(1+\beta^2(T-i))}{\sigma^2}z(1+\beta)^{i-1}}_{\text{(B)}} +$$

$$+ \underbrace{\sum_{j=i+1}^{T} \frac{\beta^2(T-j)-\beta}{\sigma^2}\beta z(1+\beta)^{j-1}}_{\text{(C)}}$$

Combine the Terms (E1) + (B)

Inside the big bracket, we see two terms that are multiples of $z(1+\beta)^{i-1}$:

$$\text{(E1)} = \frac{\beta^2(T-i)-\beta}{\sigma^2}z(1+\beta)^{i-1}$$

$$\text{(B)} = \frac{\beta(1+\beta^2(T-i))}{\sigma^2}z(1+\beta)^{i-1}$$

Add them:

$$\text{(E1)} + \text{(B)} = \frac{z(1+\beta)^{i-1}}{\sigma^2}\left[\beta^2(T-i) - \beta + \beta(1+\beta^2(T-i))\right]$$

Inside the bracket:

$$\beta^2(T-i) - \beta + \beta + \beta^3(T-i) = \beta^2(T-i) + \beta^3(T-i) = \beta^2(T-i)(1+\beta)$$

Hence,

$$\text{(E1)} + \text{(B)} = \frac{z(1+\beta)^{i-1}}{\sigma^2}\beta^2(T-i)(1+\beta)$$

Factor out $\beta^2(1+\beta)$:

$$\text{(E1)} + \text{(B)} = \frac{\beta^2(T-i)}{\sigma^2}z(1+\beta)^{i}$$

(We just pulled out one extra factor of $(1+\beta)$.)

Thus so far the bracket is:

$$[\text{(E1)} + \text{(B)}] + \text{(C)} = \frac{\beta^2(T-i)}{\sigma^2}z(1+\beta)^{i} + \text{(C)}$$

Now Add (C)

Recall (C) explicitly:

$$\text{(C)} = \sum_{j=i+1}^{T} \frac{\beta^2(T-j)-\beta}{\sigma^2} \beta z(1+\beta)^{j-1}$$

So,

$$\text{(E1)} + \text{(B)} + \text{(C)} = \frac{\beta^2(T-i)}{\sigma^2} z(1+\beta)^i + \beta z \sum_{j=i+1}^{T} \frac{\beta^2(T-j)-\beta}{\sigma^2}(1+\beta)^{j-1}$$

Factor out $\frac{z}{\sigma^2}$ to get:

$$\text{(E1)} + \text{(B)} + \text{(C)} = \frac{z}{\sigma^2}\left[\beta^2(T-i)(1+\beta)^i + \sum_{j=i+1}^{T} \beta(\beta^2(T-j)-\beta)(1+\beta)^{j-1}\right]$$

4.7 Model and Notation

We have one risky asset $\{X_t\}_{t=0,1,\ldots,T}$ satisfying

$$X_t = (1+\beta)X_{t-1} + \sigma\epsilon_t, \quad X_0 = 0, \quad \epsilon_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

Hence

$$\Delta X_t = X_t - X_{t-1} = \beta X_{t-1} + \sigma\epsilon_t$$

We fix a finite horizon T.

4.8 The First Paper's Formula Simplifies When $\mu = 0$

When $\mu_j = 0$ for all j, the term

$$\sum_{j=1}^{T} \Lambda_{i,j}\mu_j$$

vanishes. Therefore, Corollary 1.3 (or Theorem 1.1) becomes

$$\hat{\gamma}_i = -\sum_{j=1}^{i-1} \Lambda_{i,j}\Delta X_j, \quad i = 1, \ldots, T$$

In other words:

$$\boxed{\hat{\gamma}_i = -\sum_{j=1}^{i-1} \Lambda_{i,j}\left[\beta X_{j-1} + \sigma\epsilon_j\right].}$$

4.9 Covariance and Precision Matrix for the AR(1) Increments

Because $X_0 = 0$ does not affect the increment-covariance structure (the increments ΔX_t still come from $\beta X_{t-1} + \sigma\epsilon_t$ with $\epsilon_t \sim \mathcal{N}(0,1)$), the covariance matrix Σ of $\Delta\mathbf{X} = (\Delta X_1, \ldots, \Delta X_T)$ is exactly the same as in the general AR(1) case. We recall:

1. $\Delta\mathbf{X} = L\epsilon$ for a certain lower-triangular L. 2. $\Sigma = LL^\top$. 3. $\Lambda = \Sigma^{-1} = (L^\top)^{-1}L^{-1} = M^\top M$, where $M = L^{-1}$ is also lower triangular.

$$\Lambda_{i,j} = (\Sigma^{-1})_{i,j} = \begin{cases} \frac{1+\beta^2(T-i)}{\sigma^2}, & i = j, \\ \frac{\beta^2(T-\max(i,j))-\beta}{\sigma^2}, & i \neq j. \end{cases}$$

4.10 Substituting $\Lambda_{i,j}$ into the Strategy

From 4.6.5 we can already see how the optimal strategy will simplify for $z = 0$. However, the general theorem's formula also simplifies nicely for such 0-mean processes, which we will demonstrate through this example. We have

$$\hat{\gamma}_i = -\sum_{j=1}^{i-1} \Lambda_{i,j}\left[\beta X_{j-1} + \sigma\epsilon_j\right], \quad i = 1, \ldots, T$$

Observe j in that sum always satisfies $j < i$. So $\max(i,j) = i$. Thus $\Lambda_{i,j}$ simplifies to

$$\Lambda_{i,j} = \frac{\beta^2(T-i)-\beta}{\sigma^2}, \quad \text{for } j < i$$

Hence

$$\hat{\gamma}_i = -\sum_{j=1}^{i-1} \frac{\beta^2(T-i)-\beta}{\sigma^2}\left[\beta X_{j-1} + \sigma\epsilon_j\right]$$

Factor out the constant $\frac{\beta^2(T-i)-\beta}{\sigma^2}$:

$$\hat{\gamma}_i = -\left[\frac{\beta^2(T-i)-\beta}{\sigma^2}\right]\sum_{j=1}^{i-1}\left[\beta X_{j-1} + \sigma\epsilon_j\right]$$

We can rewrite this as:

$$\boxed{\hat{\gamma}_i = -\frac{\beta^2(T-i)-\beta}{\sigma^2}\left[\beta\sum_{j=1}^{i-1} X_{j-1} + \sigma\sum_{j=1}^{i-1}\epsilon_j\right].}$$

4.10.1 Sum of X_{j-1} for $j = 1, \ldots, i-1$

$$\sum_{j=1}^{i-1} X_{j-1} = X_0 + X_1 + \cdots + X_{i-2}$$

But $X_0 = 0$. Also, each X_k is a linear combination of $\{\epsilon_m\}_{m\le k}$. Concretely,

$$X_k = \sigma \sum_{m=1}^{k} (1+\beta)^{k-m} \epsilon_m$$

So

$$\sum_{j=1}^{i-1} X_{j-1} = \sum_{k=0}^{i-2} X_k = \sum_{k=0}^{i-2} \left[\sigma \sum_{m=1}^{k} (1+\beta)^{k-m} \epsilon_m \right]$$

When $k = 0$, $X_0 = 0$. So effectively the outer sum starts from $k = 1$. We can interchange sums:

$$\sum_{k=1}^{i-2} \sum_{m=1}^{k} = \sum_{m=1}^{i-2} \sum_{k=m}^{i-2}$$

Hence

$$\sum_{j=1}^{i-1} X_{j-1} = \sigma \sum_{m=1}^{i-2} \sum_{k=m}^{i-2} (1+\beta)^{k-m} \epsilon_m$$

Now the inside sum in k from m to $(i-2)$ is a finite geometric series in $(1+\beta)$. Specifically,

$$\sum_{k=m}^{i-2} (1+\beta)^{k-m} = \sum_{r=0}^{i-2-m} (1+\beta)^r = \frac{(1+\beta)^{i-1-m} - 1}{(1+\beta) - 1} = \frac{(1+\beta)^{i-1-m} - 1}{\beta}$$

So

$$\sum_{j=1}^{i-1} X_{j-1} = \sigma \sum_{m=1}^{i-2} \left[\frac{(1+\beta)^{i-1-m} - 1}{\beta} \epsilon_m \right]$$

Factor out $\frac{\sigma}{\beta}$:

$$\sum_{j=1}^{i-1} X_{j-1} = \frac{\sigma}{\beta} \sum_{m=1}^{i-2} [(1+\beta)^{i-1-m} - 1] \epsilon_m$$

4.11 Substitute These Sums Back into $\hat{\gamma}_i$

Recall:

$$\hat{\gamma}_i = -\frac{\beta^2(T-i) - \beta}{\sigma^2} \left[\beta \sum_{j=1}^{i-1} X_{j-1} + \sigma \sum_{j=1}^{i-1} \epsilon_j \right]$$

We found:

$$\beta \sum_{j=1}^{i-1} X_{j-1} = \beta \cdot \frac{\sigma}{\beta} \sum_{m=1}^{i-2} [(1+\beta)^{i-1-m} - 1] \epsilon_m = \sigma \sum_{m=1}^{i-2} [(1+\beta)^{i-1-m} - 1] \epsilon_m$$

So inside the bracket we have:

$$\beta \sum_{j=1}^{i-1} X_{j-1} + \sigma \sum_{j=1}^{i-1} \epsilon_j = \sigma \sum_{m=1}^{i-2} [(1+\beta)^{i-1-m} - 1]\epsilon_m + \sigma \sum_{j=1}^{i-1} \epsilon_j$$

We can merge these sums:

1. The first sum has m from 1 to $(i-2)$. 2. The second sum has j from 1 to $(i-1)$. Rename $j \to m$ in the second sum to see both are sums in ϵ_m. Then

$$\sigma \sum_{m=1}^{i-1} \epsilon_m = \sigma \left(\epsilon_1 + \epsilon_2 + \cdots + \epsilon_{i-1}\right)$$

Combine with the first sum:

$$[(1+\beta)^{i-1-m} - 1] + 1 = (1+\beta)^{i-1-m}$$

Hence

$$\beta \sum_{j=1}^{i-1} X_{j-1} + \sigma \sum_{j=1}^{i-1} \epsilon_j = \sigma \sum_{m=1}^{i-1} (1+\beta)^{i-1-m} \epsilon_m$$

(We used the fact that when $m = i-1$, $(1+\beta)^{i-1-(i-1)} = (1+\beta)^0 = 1$.)

But now notice:

$$\sigma \sum_{m=1}^{i-1} (1+\beta)^{i-1-m} \epsilon_m = X_{i-1},$$

because

$$X_{i-1} = \sigma \sum_{m=1}^{i-1} (1+\beta)^{i-1-m} \epsilon_m \quad \text{(see §5.1)}$$

Therefore the bracket becomes simply X_{i-1}. In short:

$$\beta \sum_{j=1}^{i-1} X_{j-1} + \sigma \sum_{j=1}^{i-1} \epsilon_j = X_{i-1}$$

4.12 Final Formula for $\hat{\gamma}_i$

We now plug this result back:

$$\hat{\gamma}_i = -\frac{\beta^2(T-i) - \beta}{\sigma^2} \underbrace{\left[\beta \sum_{j=1}^{i-1} X_{j-1} + \sigma \sum_{j=1}^{i-1} \epsilon_j\right]}_{=X_{i-1}} = -\frac{\beta^2(T-i) - \beta}{\sigma^2} X_{i-1}$$

We can rearrange the factor $\beta^2(T-i) - \beta$. Notice:

$$\beta^2(T-i) - \beta = \beta\left[\beta(T-i) - 1\right] = -\beta\left[1 - \beta(T-i)\right]$$

Hence

$$-\left[\beta^2(T-i)-\beta\right]=\beta\left[1-(T-i)\beta\right]$$

So:

$$\hat{\gamma}_i=-\frac{\left[\beta^2(T-i)-\beta\right]}{\sigma^2}X_{i-1}=\frac{\beta}{\sigma^2}\left[1-(T-i)\beta\right]X_{i-1}$$

Thus:

$$\boxed{\hat{\gamma}_i=\frac{\beta}{\sigma^2}\left[1-(T-i)\beta\right]X_{i-1}.}$$

This matches exactly the second paper's Theorem 2.1 for the fully informed (no-delay) investor.

Chapter 5

Conclusion and Outlook

This thesis set out to show the concrete manifestation of general theory explicitly on models. Starting from the abstract decomposition result of Dolinsky and Zuk (2023), to two case studies which illustrated the analytical and practical aspects of this framework. Together, these examples demonstrate both the power and the limitations of such models. As we saw, coming to an explicit solution can be rather challenging, which is why numerical approximations are something to look out for.

Looking ahead, at least three directions appear particularly promising. First, extending the delayed-information paradigm to a multi-asset settings would test the scalability of the matrix-decomposition approach. Second, relaxing the Gaussian assumption could reveal whether similar "banded + sparse" structures survive in more realistic markets. Finally, further sophisticating the assumptions, for instance the delay itself also being a random variable.

Bibliography

Wayne W. Barrett and Philip J. Feinsilver. Inverses of banded matrices. Linear Algebra and its Applications, 41:111–130, 1981. doi: 10.1016/0024-3795(81)90092-6.

Sándor Deák and Miklós Rásonyi. An explicit solution for optimal investment problems with autoregressive prices and exponential utility. Applicationes Mathematicae, 42(4):379–401, 2015. doi: 10.4064/am2267-12-2015.

Yan Dolinsky and Or Zuk. Short communication: Exponential utility maximization in a discrete time gaussian framework. SIAM Journal on Financial Mathematics, 14(3):SC31–SC41, 2023. doi: 10.1137/23M1576074.

Hans Föllmer and Alexander Schied. Stochastic Finance: An Introduction in Discrete Time. De Gruyter, Berlin / Boston, 4 edition, 2016. doi: 10.1515/9783110463453.

George T. Gilbert. Positive definite matrices and Sylvester's criterion. The American Mathematical Monthly, 98(1):44–46, 1991. doi: 10.1080/00029890.1991.11995702.

MI. nyilatkozat

Alulírott Terényi Tamás Álmos, nyilatkozom, hogy szakdolgozatom elkészítése során az alább felsorolt feladatok elvégzésére a megadott MI alapú eszközöket alkalmaztam:

Feladat	**Felhasznált eszköz**	**Felhasználás helye**
Szövegvázlat készítés	**ChatGPT o1**	Bevezetés
Nyelvhelyesség ellenőrzése/javítása/ /átfogalmazások	**ChatGPT o1**	Teljes dolgozat
LaTeX formázási tanács	**ChatGPT o1**	Teljes dolgozat
Kód hibakeresés	**ChatGPT o1**	SymPy kód

A felsoroltakon túl más MI alapú eszközt nem használtam.