Some Notes from the Book: Empirical Market Microstructure: The Institutations, Economics, and Econometrics of Securities Trading by Joel Hasbrouck

John L. Weatherwax*

October 16, 2003

Introduction

Here you'll find some notes that I wrote up as I worked through this excellent book. Much of my motivation for writing these notes was to develop a document where one could directly "read" the mathematical derivations. Too often (I feel) textbooks make jumps between equations and it can be difficult to understand the resulting flow without spending a significant amount of time deriving the given statements. With this document, hopefully one will be able to follow the more detailed and simple steps presented here to verify many of the mathematical statements made in the book. If there is any problem with this approach is that some people may find it onerous to read mathematical statement they deem to be trivial. On this matter, I tried to error on the side of completeness rather than on the side of brevity. The goal in mind was always to end with a document which could be "read" without having to do any external calculations to obtain/verify the given expressions.

I've worked hard to make these notes as good as I can, but I have no illusions that they are perfect. If you feel that that there is a better way to accomplish or explain an exercise or derivation presented in these notes; or that one or more of the explanations is unclear, incomplete, or misleading, please tell me. If you find an error of any kind – technical, grammatical, typographical, whatever – please tell me that, too. I'll gladly add to the acknowledgments in later printings the name of the first person to bring each problem to my attention.

*wax@alum.mit.edu

All comments (no matter how small) are much appreciated. In fact, if you find these notes useful I would appreciate a contribution in the form of a solution to a problem that I did not work, a mathematical derivation of a statement or comment made in the book that was unclear, a piece of code that implements one of the algorithms discussed, or a correction to a typo (spelling, grammar, etc). Sort of a "take a penny, leave a penny" type of approach. Remember: pay it forward.

Chapter 3 (The Roll Model of Trade Prices)

Notes on the text

Notes on the Roll Model

In the Roll model we assume that m_t is our midquote price often considered to be the individual stocks "fair price" which has an arithmetic random walk model for its dynamics given by

$$m_t = m_{t-1} + u_t \,. \tag{1}$$

If we assume that the (half) bid-ask spread is *constant* then we define

$$2c = a_t - b_t \,, \tag{2}$$

where a_t and b_t are the time dependent ask and bid prices respectively. With this definition c is the "half spread" and then if trades take place on the bid for a customer sell and on the ask for a buy the *trade* prices p_t are given by

$$p_t = m_t + cq_t \,, \tag{3}$$

where q_t is $+1$ when a customer is buying (from the ask) and or -1 when a customer is selling (from the bid). Assuming u_t and q_t are uncorrelated, independent, zero mean, and identically distributed we can compute second order statistics of the *change* in trade prices $\Delta p_t \equiv p_t - p_{t-1}$. Using the above we can compute

$$\begin{aligned}
\Delta p_t = p_t - p_{t-1} &= m_t + cq_t - m_{t-1} - cq_{t-1} \\
&= m_{t-1} + u_t + cq_t - m_{t-1} - cq_{t-1} \\
&= u_t + cq_t - cq_{t-1} \,. \qquad (4)
\end{aligned}$$

From this the expectation is

$$E[\Delta p_t] = E[u_t] + cE[q_t - q_{t-1}] = 0 + cE[q_t] + cE[q_{t-1}] = 0 \,.$$

The variance of Δp_t using Equation 4 is then given by

$$\begin{aligned}
\gamma_0 = \text{Var}(\Delta p_t) &= E[\Delta {p_t}^2] = E[(u_t + c(q_t - q_{t-1}))^2] \\
&= E[u_t^2 + 2cu_t(q_t - q_{t-1}) + c^2(q_t - q_{t-1})^2] \qquad (5) \\
&= E[u_t^2] + c^2E[(q_t - q_{t-1})^2] = E[u_t^2] + c^2E[q_t^2 - 2q_tq_{t-1} + q_{t-1}^2] \qquad (6) \\
&= E[u_t^2] + c^2E[q_t^2] + c^2E[q_{t-1}^2] \qquad (7) \\
&= \sigma_u^2 + 2c^2 \,, \qquad (8)
\end{aligned}$$

Note that in going from line 5 to line 6 we are using the assumption that $E[u_tq_{t-k}] = 0$ for all k. In going from line 6 to line 7 we are using the assumption that $E[q_tq_{t-k}] = 0$ for all

k. In the last line we have used $E[q_t^2] = 1$. Next we compute the lag one covariance of the change in price Δp_t. We have

$$
\begin{aligned}
\gamma_1 &= \mathrm{Cov}(\Delta p_t, \Delta p_{t-1}) = E[\Delta p_{t-1} \Delta p_t] \\
&= E[(u_t + c(q_t - q_{t-1}))(u_{t-1} + c(q_{t-1} - q_{t-2}))] \\
&= E[u_t u_{t-1} + c u_t (q_{t-1} - q_{t-2}) + c u_{t-1}(q_t - q_{t-1}) + c^2 (q_t - q_{t-1})(q_{t-1} - q_{t-2})] \quad (9) \\
&= 0 + 0 + 0 + c^2 E[(q_t - q_{t-1})(q_{t-1} - q_{t-2})] = c^2 E[q_t q_{t-1} - q_t q_{t-2} - q_{t-1}^2 + q_{t-1} q_{t-2}] \quad (10) \\
&= -c^2 E[q_{t-1}^2] = -c^2 \,. \quad (11)
\end{aligned}
$$

the same expression as in the book. Using these two equation via measuring γ_0 and γ_1 we can estimate c and σ_u^2. This is done with the following

$$
\begin{aligned}
c &= \sqrt{-\gamma_1} \quad (12) \\
\sigma_u^2 &= \gamma_0 + 2\gamma_1 \,. \quad (13)
\end{aligned}
$$

From the discussion in the book we can conclude that the given variables we have just estimated are of value because

- c is another measure of the bid ask spread or the uncertainty around the midquote. We can think about this as representing the *measurement noise* of the fair price.
- σ_u^2 is the variance of the midquote m_t price dynamics. We can think about this as representing the *process noise* of the fair price.

There are several problems (or directions for further modeling) with the above approach.

- The half spread c is not constant but is in fact time dependent.
- The trade direction indicator q_t are correlated (buy tend to follow buys and sells tend to follow sells). Thus $E[q_t q_{t-1}] \neq 0$. See Page 7 for some of the mathematics in this case.
- The trade direction indicator q_t and the movement in the midquote u_t are also correlated for much of the same reason as the previous comment. See Page 7 for some of the mathematics in this case.

There is an implementation of a sample based autocovariance estimator in the python class `Autocovariance.py`. There is an implementation of estimate of c and σ_u^2 from Roll's model in the python class `RollModel.py`.

Chapter 4 (Univariate Time-Series Analysis)

Notes on the text

Notes on moving average models

We will estimate $\gamma_0 \equiv E[\Delta p_t^2]$ and $\gamma_1 \equiv E[\Delta p_t \Delta p_{t-1}]$ from a time series of trade data p_t. If we desire to fit a MA(1) model to Δp_t we first recall that an MA(1) model for Δp_t has the following form

$$\Delta p_t = \varepsilon_t + \theta \varepsilon_{t-1} \,. \tag{14}$$

Such a model has γ_0 and γ_1 related to its parameter θ and the variance of the noise term ε_t we have

$$\gamma_0 = (1+\theta^2)\sigma_\varepsilon^2 \tag{15}$$

$$\gamma_1 = \theta \sigma_\varepsilon^2 \,. \tag{16}$$

Given that we "know" γ_0 and γ_1 by using our time series of trades we can solve for the parameters θ and σ_ε^2. If we divide the second equation by the first we get

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\theta}{1+\theta^2} \,,$$

where I have used the definition of the lag-one autocorrelation ρ_1. We can write this as a quadratic equation in θ_1 in terms of ρ_1. Putting this equation in the standard form for quadratic equation we get

$$\theta^2 - \frac{1}{\rho_1}\theta + 1 = 0 \,.$$

Solving for θ in the above we get

$$\theta = \frac{\frac{1}{\rho_1} \pm \sqrt{\frac{1}{\rho_1} - 4(1)}}{2} = \frac{1 \pm \sqrt{1 - 4\rho_1^2}}{2\rho_1} = \frac{\gamma_0 \pm \sqrt{\gamma_0^2 - 4\gamma_1^2}}{2\gamma_1} \tag{17}$$

With this expression we compute $\sigma_\varepsilon^2 = \frac{\gamma_1}{\theta}$ and find

$$\sigma_\varepsilon^2 = \frac{2\gamma_1^2}{\gamma_0 \pm \sqrt{\gamma_0^2 - 4\gamma_1^2}} \times \left(\frac{\gamma_0 \mp \sqrt{\gamma_0^2 - 4\gamma_1^2}}{\gamma_0 \mp \sqrt{\gamma_0^2 - 4\gamma_1^2}} \right) = \frac{\gamma_0 \pm \sqrt{\gamma_0^2 - 4\gamma_1^2}}{2} \,, \tag{18}$$

when we simplify.

Notes on autoregressive models

Assuming a MA(1) model for Δp_t given by Equation 14 we can write this as an AR(1) model by first solving for ε_t to get $\varepsilon_t = \Delta p_t - \theta \varepsilon_{t-1}$ and then using this expression to recursively

q_t	q_{t+1}	q_{t+2}	Unnormalized Probability	$v_{t+1} = q_{t+1} - \phi q_t$	$v_{t+2} = q_{t+2} - \phi q_{t+1}$
+1	+1	+1	α^2	$1 - \phi$	$1 - \phi$
+1	+1	-1	$\alpha(1-\alpha)$	$1 - \phi$	$-1 - \phi$
+1	-1	+1	$(1-\alpha)^2$	$-1 - \phi$	$1 + \phi$
+1	-1	-1	$\alpha(1-\alpha)$	$-1 - \phi$	$-1 + \phi$
-1	+1	+1	$(1-\alpha)\alpha$	$1 + \phi$	$1 - \phi$
-1	+1	-1	$(1-\alpha)^2$	$1 + \phi$	$-1 - \phi$
-1	-1	+1	$\alpha(1-\alpha)$	$-1 + \phi$	$1 + \phi$
-1	-1	-1	α^2	$-1 + \phi$	$-1 + \phi$

Table 1: The possible values for the variables: q_t, q_{t+1}, and q_{t+2}

replace ε_{t-1} in $\Delta p_t = \varepsilon_t + \theta\varepsilon_{t-1}$. For example, we have

$$\begin{aligned}
\Delta p_t &= \varepsilon_t + \theta\varepsilon_{t-1} \\
&= \varepsilon_t + \theta(\Delta p_{t-1} - \theta\varepsilon_{t-2}) = \varepsilon_t + \theta\Delta p_{t-1} - \theta^2\varepsilon_{t-2} \\
&= \varepsilon_t + \theta\Delta p_{t-1} - \theta^2(\Delta p_{t-2} - \theta\varepsilon_{t-3}) = \varepsilon_t + \theta\Delta p_{t-1} - \theta^2\Delta p_{t-2} + \theta^3\varepsilon_{t-3} \\
&\vdots \\
&= \varepsilon_t - \sum_{k=1}^{N}(-1)^k\theta^k\Delta p_{t-k} - (-1)^{N+1}\theta^{N+1}\varepsilon_{t-(N+1)} \quad \text{for} \quad N \geq 1\,.
\end{aligned}$$

When we take the limit $N \to \infty$ we get the AR(1) representation of our MA(1) model.

Exercise 4.1 (correlated trades in the Roll model)

Part (a): First consider all of the possible values that the three samples q_t, q_{t+1}, and q_{t+2} can take. These are tabulated in Table 1 along with the unnormalized probability of each of these transitions and the value of expressions $v_{t+1} = q_{t+1} - \phi q_t$. If we sum the values in the unnormalized probabilty column above we get the numerical value of 2. Thus to convert everything to a true normalized probability (all events sum to 1) we need to divide each unormalized probability by 2. Using the above probabilities we can compute the expression $E[v_{t+1}v_{t+2}]$. Which when we simplify is given by

$$-4\alpha^2\phi - (1+\phi)^2 + 2\alpha(1+\phi)^2\,.$$

See the Mathematical file `chap_4_prob_1.nb`. Since we must have this equal to 0 we can solve for ϕ in terms of α and get

$$\phi = -\frac{1}{1-2\alpha} \quad \text{or} \quad \phi = -1 + 2\alpha\,.$$

Since $0 < \alpha < 1$ we can write this as $-1 < -1 + 2\alpha < +1$ which is the invertible region for ϕ and thus we must take the second solution.

Part (b): Note that using the probabilities above we can show that $E[v_{t+1}] = E[v_{t+2}] = 0$. Using the expression above we can calculate $E[v_{t+1}^2]$ and $E[v_{t+2}^2]$ which must both be equal to σ_v^2.

Part (c): In the above Mathematica file we can compute $E[v_{t+1}v_{t+2}^3]$ (note the 3 on the expression for v_{t+2} my version of the book has a two) and get

$$-32(-1+\alpha)^2\alpha^2(-1+2\alpha)\,.$$

Exercise 4.2 (autocorrelated trades)

Part (a): To evaluate $\text{Var}(\Delta p_t)$ when the trade indicators q_t are correlated we start with Equation 6 but then replace $E[q_t q_{t-1}]$ with

$$E[q_t q_{t-1}] = \sqrt{\text{Var}(q_t)}\sqrt{\text{Var}(q_t)}\text{Corr}(q_t, q_{t-1}) = \text{Corr}(q_t, q_{t-1}) = \rho\,.$$

When we do this we get

$$\text{Var}(\Delta p_t) = \sigma_u^2 + 2c^2 - 2c^2\rho = \sigma_u^2 + 2c^2(1-\rho)\,,$$

as we were to show. To evaluate $\text{Cov}(\Delta p_t, \Delta p_{t-1})$ we can start with Equation 10 from which we get

$$\begin{aligned}\gamma_1 &= c^2 E[q_t q_{t-1} - q_t q_{t-2} - q_{t-1}^2 + q_{t-1}q_{t-2}] \\ &= c^2(\rho - 0 - 1 + \rho) = -c^2(1-2\rho)\,,\end{aligned}$$

as we were to show. To evaluate $\text{Cov}(\Delta p_t, \Delta p_{t-2})$ we have

$$\begin{aligned}\text{Cov}(\Delta p_t, \Delta p_{t-2}) &= E[\Delta p_t \Delta p_{t-2}] \\ &= E[(u_t + c(q_t - q_{t-1}))(u_{t-1} + c(q_{t-2} - q_{t-3}))] \\ &= E[u u_{t-2}] + 0 + 0 + c^2 E[(q_t - q_{t-1})(q_{t-2} - q_{t-3})] \\ &= 0 - c^2 E[q_{t-1}q_{t-2}] = -c^2\rho\,.\end{aligned}$$

We have $\text{Cov}(\Delta p_t, \Delta p_{t-k}) = 0$ for $k \geq 3$.

Part (b): If our true process had autocorrelated trades and we used the default Roll model to estimate the half spread c via Equation 12 we would say

$$\hat{c} = \sqrt{-\gamma_1} = \sqrt{c^2(1-2\rho)} = c\sqrt{1-2\rho}\,.$$

Since we are assuming that $0 < \rho < 1$ we can manipulate this to show that $|1-2\rho| < 1$ and thus from the definition of $\hat{c}$ above we have

$$|\hat{c}^2| = |c^2(1-2\rho)| \leq |c^2|\,,$$

showing that $\hat{c}$ *underestimates* the true value of c.

Exercise 4.3 (correlated trades and price direction)

Part (a): To evaluate $\text{Var}(\Delta p_t)$ when the trade indicators q_t are correlated with the fair noise n_t we start with Equation 5 but then replace $E[u_t q_t] = \sigma_u \rho$. When we do that we get

$$\text{Var}(\Delta p_t) = \sigma_u^2 + 2c\sigma_u\rho + c^2(1 + 1 - 2E[q_t q_{t-1}]) = 2c^2 + 2c\sigma_u\rho + \sigma_u^2\,.$$

To evaluate $\text{Cov}(\Delta p_t, \Delta p_{t-1})$ we can start with Equation 9 and get

$$\text{Cov}(\Delta p_t, \Delta p_{t-1}) = 0 + 0 + 0 - c\sigma_u \rho + c^2(-1) = -c(c + \rho\sigma_u)\,.$$

We have $\text{Cov}(\Delta p_t, \Delta p_{t-k}) = 0$ for $k \geq 2$.

Part (b): If our true process had correlated price movement and trades and we incorrectly used the default Roll model to estimate the half spread c via Equation 12 we would think

$$\hat{c} = \sqrt{-\gamma_1} = \sqrt{c^2 + c\rho\sigma_u} > c\,,$$

showing that in this case we would over estimate our half spread.

Chapter 5 (Sequential Trade Models)

Notes on the text

Notes on a simple sequential trade model

From the diagram given in the book we have that the probability of a buy is given by

$$\begin{aligned}\Pr(\text{Buy}) &= 0 + \frac{1}{2}(1-\mu)\delta + \mu(1-\delta) + \frac{1}{2}(1-\mu)(1-\delta) \\ &= \frac{1}{2}(1+\mu(1-2\delta))\,, \end{aligned} \quad (19)$$

when we simplify. In the same way for sells we have

$$\begin{aligned}\Pr(\text{Sell}) &= \mu\delta + \frac{1}{2}(1-\mu)\delta + \frac{1}{2}(1-\mu)(1-\delta) \\ &= \frac{1}{2}(1-\mu(1-2\delta))\,, \end{aligned} \quad (20)$$

when we simplify. Note from the above expressions that $\Pr(\text{Buy}) + \Pr(\text{Sell}) = 1$ as they should. If we are in the case where $\Pr(\text{Buy}) = \Pr(\text{Sell})$ then we must have (after canceling the common $\frac{1}{2}$ on both sides)

$$1 + \mu(1-2\delta) = 1 - \mu(1-2\delta)\,,$$

or

$$2\mu(1-2\delta) = 0\,.$$

This means that $\mu = 0$ (no "informed" traders) or $1 - 2\delta = 0$ or $\delta = \frac{1}{2}$ (no directional movement of the price V). We will now estimate our changes in belief in the fair value of the tradable given that a trade (buy/sell) has taken place. Note that to do this we only need to compute the probability that the fair price is *less than* V given the trade direction. That is, we only need to evaluate the expressions $\Pr(\underline{V}|\text{Buy})$ and $\Pr(\underline{V}|\text{Sell})$. The reason for this is that once we have these two expressions, to calculate the probability that the fair *is greater* than V given the trade direction we simply use

$$\begin{aligned}\Pr(\overline{V}|\text{Buy}) &= 1 - \Pr(\underline{V}|\text{Buy}) \\ \Pr(\overline{V}|\text{Sell}) &= 1 - \Pr(\underline{V}|\text{Sell})\,.\end{aligned}$$

A buy trade occurs

We now compute the dealers updated belief that the true price is $\underline{V}$ (less than the current midquote V) based on the observation that the last trade was a *buy*. Using the diagram in the book we have

$$\Pr(\underline{V}|\text{Buy}) \equiv \delta_1(\text{Buy}) = \frac{\Pr(\underline{V},\text{Buy})}{\Pr(\text{Buy})} = \frac{\frac{1}{2}(1-\mu)\delta}{\frac{1}{2}(1+\mu(1-2\delta))} = \frac{\delta(1-\mu)}{1+\mu(1-2\delta)}\,. \quad (21)$$

Note that δ is the a prior probability that the stock moves "down". The notation $\delta_1(\text{Buy})$ is the new probability that the stock moves "down" given that one buy trade has occurred. Given the above expression we can compute $\Pr(\overline{V}|\text{Buy})$ using

$$\begin{aligned}\Pr(\overline{V}|\text{Buy}) = 1 - \delta_1(\text{Buy}) &= 1 - \frac{\delta(1-\mu)}{1+\mu(1-2\delta)} \\ &= \frac{1-\delta+\mu-\mu\delta}{1+\mu(1-2\delta)} = \frac{(1-\delta)(1+\mu)}{1+\mu(1-2\delta)}. \end{aligned} \quad (22)$$

We can show that as μ increases (we have more informed traders) we expect that $\delta_1(\text{Buy})$ to decrease since for each buy the dealer observes is less likely to have come from an uninformed trader. Each buy that comes from an informed trader is expected to indicate that the price will move up. Taking the needed derivatives we have

$$\begin{aligned}\frac{\partial \delta_1(\text{Buy})}{\partial \mu} &= -\frac{\delta}{1+\mu(1-2\delta)} - \frac{\delta(1-\mu)(1-2\delta)}{(1+\mu(1-2\delta))^2} \\ &= -\frac{2\delta(1-\delta)}{(1+\mu(1-2\delta))^2},\end{aligned}$$

when we simplify. As $0 < \delta < 1$ we also have $0 < 1-\delta < 1$ and the numerator above is positive which means that the entire expression for the derivative is negative. If the market maker sells at the ask A, then his profit Π is $\Pi = A - V$ since originally the security he held was worth V. His expected profit when he sells for A and someone buys is given by

$$\begin{aligned}E[\Pi|\text{Buy}] &= A - E[V|\text{Buy}] \\ &= A - [\Pr(\underline{V}|\text{Buy})\underline{V} + (1-\Pr(\underline{V}|\text{Buy}))\overline{V}] \\ &= A - [\delta_1(\text{Buy})\underline{V} + (1-\delta_1(\text{Buy}))\overline{V}].\end{aligned}$$

If we assume that competition drives the expected profit $E[\Pi|\text{Buy}]$ to zero (otherwise everyone would sell at the ask) then we have from our expressions for $\delta_1(\text{Buy})$ via Equation 21 and 22 that

$$\begin{aligned}A = E[V|\text{Buy}] &= \frac{\delta(1-\mu)}{1+\mu(1-2\delta)}\underline{V} + \frac{(1-\delta)(1+\mu)}{1+\mu(1-2\delta)}\overline{V} \\ &= \frac{\underline{V}(1-\mu)\delta + \overline{V}(1-\delta)(1+\mu)}{1+\mu(1-2\delta)}.\end{aligned} \quad (23)$$

In words, the equation $A = E[V|\text{Buy}]$ states that the dealers ask is the expected value given that someone is going to buy at that price. In other words, the expected price at which when one buys at $\overline{V}$ there is no more edge in the trade. We now consider the case where a sell trade occurs.

A sell trade occurs

For bids where people sell to the dealer we have

$$\begin{aligned}\Pr(\underline{V}|\text{Sell}) \equiv \delta_1(\text{Sell}) = \frac{\Pr(\underline{V},\text{Sell})}{\Pr(\text{Sell})} &= \frac{\mu\delta + \frac{1}{2}(1-\mu)\delta}{\frac{1}{2}(1-\mu(1-2\delta))} \\ &= \frac{\delta(1+\mu)}{1-\mu(1-2\delta)}.\end{aligned} \quad (24)$$

In the same way as before we compute

$$
\begin{aligned}
\Pr(\overline{V}|\text{Sell}) = 1 - \delta_1(\text{Sell}) &= \frac{1 - \mu(1-2\delta) - \delta - \delta\mu}{1 - (1-2\delta)\mu} \\
&= \frac{(1-\mu)(1-\delta)}{1 - \mu(1-2\delta)} \,. \qquad (25)
\end{aligned}
$$

when we simplify. Since we expect that $\underline{V}$ is more likely when a sell occurs relative to when a buy occurs we expect that our model should show that $\Pr(\underline{V}|\text{Sell}) > \Pr(\underline{V}|\text{Buy})$. We will now show this fact which involves some inequality manipulations. From the expressions just computed we can evaluate the ratio

$$
\frac{\Pr(\underline{V}|\text{Sell})}{\Pr(\underline{V}|\text{Buy})} = \left(\frac{1+\mu}{1-\mu}\right)\left(\frac{1+\mu(1-2\delta)}{1-\mu(1-2\delta)}\right) .
$$

Since $0 < \delta < 1$ we have that $-2 < -2\delta < 0$ and $-1 < 1-2\delta < +1$ or that the variable $1-2\delta$ is such that $|1-2\delta| < 1$. Since μ is positive we can write this as

$$
\mu|1-2\delta| < \mu \,,
$$

or

$$
-\mu < \mu(1-2\delta) < +\mu \quad \text{and} \quad -\mu < -\mu(1-2\delta) < +\mu \,.
$$

when we multiply by a negative one. Adding one to each of these inequalities we get

$$
1-\mu < 1+\mu(1-2\delta) < 1+\mu \quad \text{and} \quad 1-\mu < 1-\mu(1-2\delta) < 1+\mu \,.
$$

Using these expressions we find a lower bound on the ratio given by

$$
\frac{1+\mu(1-2\delta)}{1-\mu(1-2\delta)} > \frac{1-\mu}{1+\mu} \,.
$$

This means that

$$
\frac{\Pr(\underline{V}|\text{Sell})}{\Pr(\underline{V}|\text{Buy})} > \left(\frac{1+\mu}{1-\mu}\right)\left(\frac{1-\mu}{1+\mu}\right) = 1 \,,
$$

as we were to show. We expect the more informed traders there are (i.e. the larger the value of μ is) that more from each person that sells to us is an indication that the stock is going to go down and that its value is more likely $\underline{V}$. We find

$$
\begin{aligned}
\frac{\partial \delta_1(\text{Sell})}{\partial \mu} &= \frac{\delta}{1-\mu(1-2\delta)} + \frac{\delta(1+\mu)(1-2\delta)}{(1-\mu(1-2\delta))^2} \\
&= \frac{2\delta(1-\delta)}{(1-\mu(1-2\delta))} \,,
\end{aligned}
$$

when we simplify. Since this expression is positive we have the requested expression. By considering the profit to the dealer when he buys (the market buys at the bid B) we can show that

$$
\begin{aligned}
B = E[V|\text{Sell}] &= \Pr(\underline{V}|\text{Sell})\underline{V} + (1 - \Pr(\underline{V}|\text{Sell}))\overline{V} \\
&= \frac{\delta(1+\mu)}{1-\mu(1-2\delta)}\underline{V} + \frac{(1-\delta)(1-\mu)}{1-\mu(1-2\delta)}\overline{V} \\
&= \frac{\underline{V}(1+\mu)\delta + \overline{V}(1-\mu)(1-\delta)}{1-\mu(1-2\delta)} \,. \qquad (26)
\end{aligned}
$$

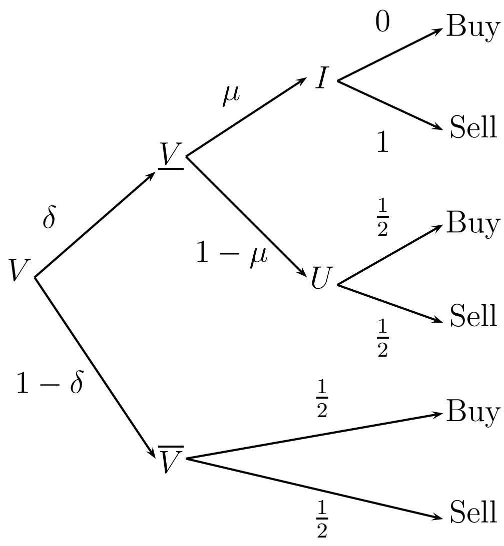

Table 2: The possible outcomes when informed trading happens only in the "low" state $\underline{V}$

The spread in dollars is then predicted to be

$$A - B = \frac{4\mu[\delta^2(\overline{V} - \underline{V}) + \overline{V} - 2\delta\overline{V}]}{1 - (1 - 2\delta)^2\mu^2} = \frac{4\mu[\delta^2\overline{V} - 2\delta\overline{V} + \overline{V} - \delta^2\underline{V}]}{1 - (1 - 2\delta)^2\mu^2}$$
$$= \frac{4\mu[(\delta - 1)^2\overline{V} - \delta^2\underline{V}]}{1 - (1 - 2\delta)^2\mu^2} . \quad (27)$$

This we verified with the Mathematica notebook `chap_5_bid_ask_spread.nb` and is somewhat different than the one given in the book. If $\delta = 1/2$ then the above expression also gives $A - B = (\overline{V} - \underline{V})\mu$.

Exercise 5.2 (informed trading only in the low $\underline{V}$ state)

We diagram the possible transitions in the case specified in Figure 2. From the figure we see that the unconditional probabilities of a buy and a sell order are

$$\Pr(\text{Buy}) = \frac{1}{2}(1 - \delta) + \frac{1}{2}\delta(1 - \mu) = \frac{1}{2}(1 - \delta\mu)$$
$$\Pr(\text{Sell}) = \frac{1}{2}(1 - \delta) + \frac{1}{2}\delta(1 - \mu) + \delta\mu = \frac{1}{2}(1 + \delta\mu) .$$

Note that $\Pr(\text{Buy}) + \Pr(\text{Sell}) = 1$ as it should. We can now compute the probability of $\underline{V}$ conditional on a buy or sell order arriving. We find

$$\delta_1(\text{Buy}) = \Pr(\underline{V}|\text{Buy}) = \frac{\Pr(\underline{V}, \text{Buy})}{\Pr(\text{Buy})} = \frac{0 + \frac{1}{2}\delta(1 - \mu)}{\frac{1}{2}(1 - \delta\mu)} = \frac{\delta(1 - \mu)}{1 - \delta\mu} .$$

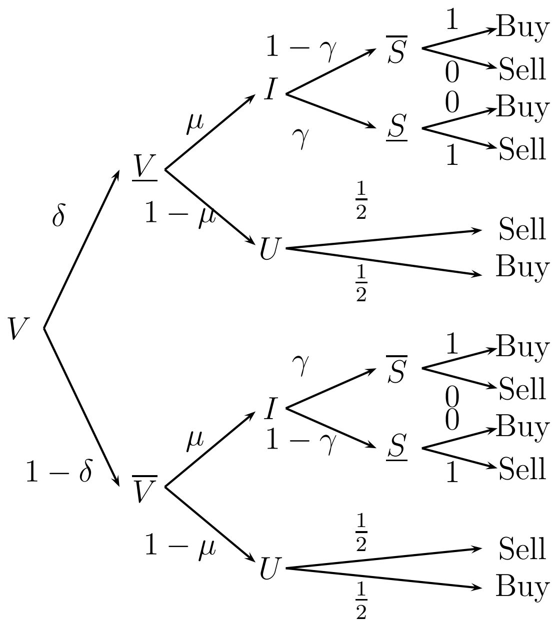

Table 3: The possible outcomes when informed traders have a signal to enter a trade.

$$\delta_1(\text{Sell}) = \Pr(\underline{V}|\text{Sell}) = \frac{\Pr(\underline{V},\text{Sell})}{\Pr(\text{Sell})} = \frac{\delta\mu + \frac{1}{2}\delta(1-\mu)}{\frac{1}{2}(1+\delta\mu)} = \frac{\delta(1+\mu)}{1+\delta\mu}.$$

Using these expressions we find for the ask A

$$\begin{aligned} A = E[V|\text{Buy}] &= \Pr(\underline{V}|\text{Buy})\underline{V} + (1 - \Pr(\underline{V}|\text{Buy}))\overline{V} \\ &= \delta_1(\text{Buy})\underline{V} + (1-\delta_1(\text{Buy}))\overline{V} \\ &= \frac{\delta(1-\mu)\underline{V} + (1-\delta)\overline{V}}{1-\delta\mu}, \end{aligned}$$

when we simplify. For the bid B we get

$$\begin{aligned} B = E[V|\text{Sell}] &= \Pr(\underline{V}|\text{Sell})\underline{V} + (1 - \Pr(\underline{V}|\text{Sell}))\overline{V} \\ &= \delta_1(\text{Sell})\underline{V} + (1-\delta_1(\text{Sell}))\overline{V} \\ &= \frac{\delta(1+\mu)\underline{V} + (1-\delta)\overline{V}}{1+\delta\mu}, \end{aligned}$$

when we simplify.

Exercise 5.3 (informed traders with a signal)

We diagram the possible transitions in the case specified in Figure 3. From the figure we see that the unconditional probabilities of a buy and a sell order are given by

$$\begin{aligned}
\Pr(\text{Buy}) &= \delta\mu(1-\gamma) + \frac{1}{2}\delta(1-\mu) + (1-\delta)\mu\gamma + \frac{1}{2}(1-\delta)(1-\mu) \\
&= \frac{1}{2} - \left(\frac{1}{2} - \delta - \gamma + 2\delta\gamma\right)\mu \\
\Pr(\text{Sell}) &= \delta\mu\gamma + \frac{1}{2}\delta(1-\mu) + (1-\delta)\mu(1-\gamma) + \frac{1}{2}(1-\delta)(1-\mu) \\
&= \frac{1}{2} + \frac{1}{2}(1-2\delta)(1-2\gamma)\mu\,.
\end{aligned}$$

Now that $\Pr(\text{Buy}) + \Pr(\text{Sell}) = 1$ as it should. Given these and from the diagram in Figure 3 we get

$$\begin{aligned}
\delta(\text{Buy}) &= \frac{\Pr(\underline{V}, \text{Buy})}{\Pr(\text{Buy})} = \frac{\delta\mu(1-\gamma) + \frac{1}{2}\delta(1-\mu)}{\Pr(\text{Buy})} \\
&= \frac{\delta(1-(2\gamma-1)\mu)}{1-(2\delta-1)(2\gamma-1)\mu}\,,
\end{aligned}$$

when we simplify.

Exercise 5.4 (offsetting trades are uninformative)

From trade sequencing we have

$$\delta_2(\text{Sell}_1, \text{Buy}_2) = \delta_2(\text{Buy}_2; \delta_1(\text{Sell}_1))\,.$$

From Equation 21 this last expression is given by

$$\frac{\delta_1(\text{Sell}_1)(1-\mu)}{1+\mu(1-2\delta_1(\text{Sell}_1))}\,.$$

Using Equation 24 to replace $\delta_1(\text{Sell}_1)$ in the above we get

$$\frac{\left(\frac{\delta(1+\mu)}{1-(1-2\delta)\mu}\right)(1-\mu)}{1+\mu\left(1-\frac{2\delta(1+\mu)}{1-(1-2\delta)\mu}\right)} = \delta\,,$$

when we simplify the fraction.

Chapter 6 (Order Flow and the Probability of Informed Trading)

Notes on the text

Notes on event uncertainty and Poisson arrivals

This is a relatively short chapter with no problems but one thing that seemed confusing to me on the first reading was the diagram given to represent event uncertainty coupled with the Poisson arrival rate for trades (buy/sells). We can reason about this diagram by understanding that depending on what external "event" happens to our stock the intensities of buying and selling will change. For example, with probability $1-\alpha$ nothing informative has happened and there is no trade for informative traders to take. In that case the Poisson intensity of buyers and sellers is *equal* and denoted by ε. On the other hand, with a probability of α, an information event has taken place. In that case, with another probability δ, this is a "down" event or $V \to \underline{V}$ and all informed traders will be selling. Thus we expect the intensity distribution of buyers vs. sellers to be have more sellers than buyers. This is denoted using Poisson intensities as the distribution $(\varepsilon, \varepsilon+\mu)$. If the event is such that $V \to \overline{V}$ then the informed traders are buying and Poisson intensity distribution is given by $(\varepsilon+\mu, \varepsilon)$.

Chapter 7 (Strategic Trade Models)

Notes on the text

Notes on the single-period Kyle model

Assume the world is divided up into informed and uniformed traders who will submit their orders and then all trades take place at a common price p which is set by the market maker (MM). The market maker has to absorb the excess liquidity when very large orders come in. We assume that the final "fair" stock price (denoted by a v) is a random variable given by a $\mathcal{N}(p_0, \Sigma_0)$ distribution. The difference between informed and uninformed traders are based on the fact that the informed traders will *know* this final price v. Since the informed traders know v they desire to trade as much stock as possible at a price that is advantageous to that final price. The informed trader will submit market orders to try to make these trades. There is a trade off between the total quantity (size) of market orders that the informed trader will submit and his impact on the price the market maker will set. For example, if the informed trader thinks that the price will go up he will submit his buy markets orders. These orders then cause the market maker increase the auction price. If the price increase too much the trade may not end up profitable.

It is in the equilibrium between the final fair price v, the sized of the order the informed trader will submit, and how the market maker adjusts his price based on that demand that determines the solution.

The informed trader submits his demand for x shares/dollars and the noise traders submits a "random" demand u for the stock given by $u \sim \mathcal{N}(0, \sigma_u^2)$. The total demand the market marker then sees is denoted by y is the sum $y = x + u$. The variables u and x are positive if traders want to buy and negative if they want to sell. The market marker seeing this total demand will set the auction price in a linear manner related to the demand as

$$p = \lambda y + \mu \,. \tag{28}$$

Here λ is a liquidity scaling parameter that specifies how the market maker will change the fill price p depending on the liquidity y observed. If there is no demand $y = 0$ then $p = \mu$ so μ is the zero demand price which would be close to the midquote of any market that traded before the trades the market maker must participate in.

Now if informed trader does not know the value of λ used by the market maker (one will be derived below) one can estimate this parameter by observing the total demand y by computing a linear regression between the response of $p - \mu$ or the difference between the open/auction price and the midquote just prior to the auction and y the total incoming imbalance messages. That is we fit the model

$$p - \mu = \lambda y + \epsilon \,,$$

using linear regression or other such method. Here ϵ is a error term. Using this model one could predict the open price given the total demand y.

As informed traders *know* the final security price v and they get filled for x against the market marker (MM) at the price p. The profit from this strategy is then known at $\pi = (v - p)x$. Using what we know since the market makers "price setting function" is assumed to be $p = \lambda y + \mu$ we can write the profit as

$$\pi = (v - \lambda y - \mu)x = (v - \mu - \lambda y)x\,.$$

Note that $v - \mu$ is the price difference between the current midmarket μ and the known final price v. In terms of our informed traders demand x we have $y = x + u$ so our profit is given by

$$\pi = (v - \mu - \lambda(x + u))x\,.$$

Since u is a random variable we will evaluate the expectation of this expression under the assumption that as an informed trader we *know* the final price v. In this case, the noise traders demand u is independent of everything else so using $E[u|v] = 0$ the expected profit is

$$E[\pi|v] = (v - \mu - \lambda x)x\,.$$

In general, for models of this type we will compute the expected profit of the informed trader given the information the informative trader has which in this case is the final price v. To maximize $E[\pi|v]$ as a function of x the informed trader would compute $\frac{dE[\pi|v]}{dx} = 0$ and solve for x. He would find

$$v - \mu - 2\lambda x = 0 \quad \text{or} \quad x = \frac{v - \mu}{2\lambda}\,. \tag{29}$$

The second order criterion (that we have found a maximum and not a minimum) is given by

$$\frac{d^2}{dx^2}E[\pi|v] = -2\lambda < 0\,.$$

This requires that $\lambda > 0$. If the market maker assumes that the informed trader acts rationally and follow the above strategy the informed trader has a submitted demand x that is linear in v (i.e. $x = \alpha + \beta v$ for some α and β) since we can write the expression in Equation 29 for x as

$$x = -\frac{\mu}{2\lambda} + \frac{1}{2\lambda}v\,.$$

Thus in the linear form $x = \alpha + \beta v$ we have

$$\alpha = -\frac{\mu}{2\lambda} \quad \text{and} \quad \beta = \frac{1}{2\lambda}\,. \tag{30}$$

The market maker will not loose money or suffer from selection bias if he can set the auction/trade price p exactly at the fair or final price v. The market maker might not know v but if he can try to compute $E[v|y]$ where y is the *total* order flow from the informed and the noise traders. The fact that the variable y can tell us something about v follows from the fact that y depends on x (the informed traders orders) and x depends on v (via the linear relationship above). Thus the variables y and v are linked and knowledge of one should help in determining the other. We will use the result on the expectation of a conditioned random variable given in the book. Writing y in terms of the final price v we have

$$y = u + \alpha + \beta v\,.$$

Since u and v are random $u \sim \mathcal{N}(0, \sigma_u^2)$ and $v \sim \mathcal{N}(p_0, \Sigma)$ we have that

$$E[y] = \alpha + \beta p_0 \tag{31}$$
$$\text{var}[y] = \sigma_u^2 + \beta^2 \Sigma_0 \tag{32}$$
$$\text{cov}(y, v) = \text{cov}(u + \alpha + \beta v, v) = 0 + \beta \text{cov}(v, v) = \beta \Sigma_0 \,. \tag{33}$$

Then using these expressions and the theorem in the book we find

$$E[v|y] = \mu_v + \frac{\sigma_{vy}}{\sigma_y^2}(y - \mu_y) \quad \text{(the definition) which in this case becomes} \tag{34}$$
$$= p_0 + \frac{\beta \Sigma_0}{\sigma_u^2 + \beta^2 \Sigma_0}(y - \alpha - \beta p_0) \,. \tag{35}$$

and

$$\text{var}(v|y) = \sigma_v^2 - \frac{\sigma_{vy}^2}{\sigma_y^2} \quad \text{(the definition) which in this case becomes} \tag{36}$$
$$= \Sigma_0 - \frac{(\beta \Sigma_0)^2}{\sigma_u^2 + \beta^2 \Sigma_0} = \frac{\sigma_u^2 \Sigma_0}{\sigma_u^2 + \beta^2 \Sigma_0} \,. \tag{37}$$

Again for the market market to not loose money he must set his price p at $E[v|y]$. Thus we require that $E[v|y]$ computed above equal match the market makers liquidity price adjustment relationship $p = \lambda y + \mu$ for all y. This gives

$$E[v|y] = p_0 - \left(\frac{\beta \Sigma_0}{\sigma_u^2 + \beta^2 \Sigma_0} \right) (\alpha + \beta p_0) + \left(\frac{\beta \Sigma_0}{\sigma_u^2 + \beta^2 \Sigma_0} \right) y = \lambda y + \mu \,.$$

Equating these two expressions when we group terms by powers of y we have that

$$\mu = p_0 - \left(\frac{\beta \Sigma_0}{\sigma_u^2 + \beta^2 \Sigma_0} \right) (\alpha + \beta p_0) = \frac{p_0 \sigma_u^2 - \alpha \beta \Sigma_0}{\sigma_u^2 + \beta^2 \Sigma_0} \,, \tag{38}$$

and

$$\lambda = \frac{\beta \Sigma_0}{\sigma_u^2 + \beta^2 \Sigma_0} \,. \tag{39}$$

Adding the expressions for the linear form of the informed traders demand $x = \alpha + \beta v$ of $\alpha = -\frac{\mu}{2\lambda}$ and $\beta = \frac{1}{2\lambda}$ with the equations 38 and 39 we have four equations for the four unknowns: α, β, μ, and λ. When we solve these four equations simultaneously in terms of the variables p_0, Σ_0 and σ_u^2 we get

$$\alpha = \mp p_0 \sqrt{\frac{\sigma_u^2}{\Sigma_0}}$$
$$\mu = p_0$$
$$\lambda = \pm \frac{1}{2} \sqrt{\frac{\Sigma_0}{\sigma_u^2}}$$
$$\beta = \pm \sqrt{\frac{\sigma_u^2}{\Sigma_0}} \,.$$

The book selects the first solution (the one with a positive value for λ since from Equation 28 we would expect that when $y > 0$ that $p > \mu$). Note this gives a *negative* value for α). See the Mathematica file `chapter_7_algebra.nb`. For the expected profit under all of these assumptions we have

$$
\begin{aligned}
E[\pi] = x(v - \mu - \lambda x) &= \left(\frac{v-\mu}{2\lambda}\right)\left(v - \mu - x\left(\frac{v-\mu}{2\lambda}\right)\right) \\
&= \frac{(v-\mu)^2}{4\lambda} = \frac{(v-p_0)^2}{2}\sqrt{\frac{\sigma_u^2}{\Sigma_0}}\,. \qquad (40)
\end{aligned}
$$

The variance in the fair price v given the incoming interest y, where we use $\beta = \sqrt{\frac{\sigma_u^2}{\Sigma_0}}$ is then given by

$$
\text{var}[v|y] = \frac{\sigma_u^2\Sigma_0}{\sigma_u^2 + \Sigma_0\left(\frac{\sigma_u^2}{\Sigma_0}\right)} = \frac{\Sigma_0}{2}\,.
$$

Exercise 7.1 (partially informed noise traders)

For this problem we will assume that $\text{Cov}(u, v) = \sigma_{uv} > 0$. In this case the expression for the informative traders profit π does not change. Namely when we use the expression Equation 28 for p we get

$$
\pi = (v - p)x = (v - \mu - \lambda(x + u))x\,.
$$

As the informed trader knows the final price v the expected profit is given by

$$
E[\pi|v] = (v - \mu - \lambda(x + E[u|v]))x\,.
$$

Since now u and v are correlated, we no longer have that $E[u|v] = 0$ as we did before. Using the expression for conditional expectations given in the book (namely Equation 34) we have that

$$
E[u|v] = E[u] + \frac{\sigma_{uv}}{\sigma_v^2}(v - E[v]) = \frac{\sigma_{uv}}{\Sigma_0}(v - p_0)\,.
$$

Note that this is independent of x, contains all known expressions, and will not change the "form" of the optimal x (except to shift it). Thus we get for the optimal order size for the informed trader

$$
x = \frac{v - \mu - \lambda E[u|v]}{2\lambda} = \frac{(v-\mu)\Sigma_0 - \lambda\sigma_{uv}(v - p_0)}{2\lambda\Sigma_0}\,.
$$

Writing this as $x = \alpha + \beta v$ we can compute α and β, that involve only known quantities. We now need to evaluate $E[v|y]$ given that the informed trader acts under his optimal strategy. With $y = u + \alpha + \beta v$ (as before) we need to compute the variance of y computed in Equation 32 and the covariance computed in Equation 33. To compute the variance we will use

$$
\text{var}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \text{var}[X_i] + 2\sum\sum_{i<j}\text{cov}[X_i, X_j]\,. \qquad (41)
$$

Using this expression we find

$$\text{var}[y] = \text{var}[u + \alpha + \beta v] = \text{var}[u + \beta v] = \sigma_u^2 + \beta^2 \Sigma_0 + 2\beta\sigma_{uv} \,.$$

With the covariance between y and v given by

$$\text{cov}(y, v) = \text{cov}(u + \alpha + \beta v, v) = \sigma_{uv} + \beta\Sigma_0 \,.$$

With these two expressions we can use Equation 34 to compute $E[v|y]$ and set the resulting expression equal to $\lambda y + \mu$. We find that

$$\begin{aligned} E[v|y] &= E[v] + \frac{\sigma_{vy}}{\sigma_y^2}(y - E[y]) \\ &= p_0 + \left(\frac{\beta\Sigma_0 + \sigma_{uv}}{\sigma_u^2 + \beta^2\Sigma_0 + 2\beta\sigma_{uv}} \right)(y - \alpha - \beta p_0) \,. \end{aligned}$$

Setting $E[v|y]$ equal to $\lambda y + \mu$ we get can solve for λ, μ α and β in terms of known parameters of the problem. This is done in the Mathematica notebook `chapter_7_algebra.nb` in the variable `wSol`.

Exercise 7.2 (the informed trader gets a signal to the fair v)

For this section of the book we assume that the market marker sets his trade/auction price at p which has a linear impact with the total order flow y as $p = \lambda y + \mu$ in the same way as before. The total order flow y is a random variable that is the sum of the informed trading request x and a random uniformed trader amount u. That is $y = x + u$ where $u \sim \mathcal{N}(0, \sigma_u^2)$. The final fair price of the security or v is a random variable with a distribution $\mathcal{N}(p_0, \Sigma_0)$. In this problem the informed trader does not know v but instead s which is derived from v as $s = v + \epsilon$ with $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. Under these conditions the informed traders profits are given as before

$$\pi = x(v - \lambda(x + u) - \mu) \,.$$

To compute the expected profit we condition on the information the informed trader knows i.e. the proxy to the true price v or s. Thus we need to evaluate

$$E[\pi|s] = x(E[v|s] - \lambda x - \mu) \,.$$

From the formula in the book for conditional expectation we have

$$\begin{aligned} E[\pi|s] &= E[\pi] + \frac{\sigma_{\pi,s}}{\sigma_s^2}(s - E[s]) \\ &= x(p_0 - \lambda x - \mu) + \frac{x\Sigma_0}{\sigma_\epsilon^2 + \Sigma_0}(s - p_0) \,. \end{aligned}$$

The informed trader wants to optimize the expected profit and loss so he solves

$$\frac{d}{dx} E[\pi|s] = p_0 - \lambda x - \mu + x(-\lambda) + \frac{\Sigma_0}{\sigma_\epsilon^2 + \Sigma_0}(s - p_0) = 0 \,,$$

for x to get

$$x = \frac{1}{2\lambda}\left(p_0 - \mu + \left(\frac{\Sigma_0}{\sigma_\epsilon^2 + \Sigma_0}\right)(s - p_0)\right) = \frac{s\Sigma_0 + \sigma_\epsilon^2 p_0 - \mu(\Sigma_0 + \sigma_\epsilon^2)}{2\lambda(\Sigma_0 + \sigma_\epsilon^2)}.$$

If we write this as $x = \alpha + \beta s$ we get for α and β

$$\alpha = \frac{\sigma_\epsilon^2 p_0 - (\sigma_\epsilon^2 + \Sigma_0)\mu}{2\lambda(\sigma_\epsilon^2 + \Sigma_0)} \tag{42}$$

$$\beta = \frac{\Sigma_0}{2\lambda(\sigma_\epsilon^2 + \Sigma_0)}. \tag{43}$$

Note that these expressions have μ and λ in them. The market market must compute $E[v|y]$ again using the expression for conditional expectation

$$E[v|y] = E[v] + \frac{\sigma_{vy}}{\sigma_y^2}(y - E[y]).$$

We now compute each of the terms needed to compute this expression

$$\begin{aligned}
\sigma_{vy} &= \text{cov}(v, y) = \text{cov}(v, x + u) = \text{cov}(v, x) = \text{cov}(v, \beta s) \\
&= \beta\text{cov}(v, s) = \beta\text{cov}(v, v + \epsilon) = \beta\Sigma_0 \\
\text{var}(y) &= \text{var}(x + u) = \text{var}(\alpha + \beta s + u) = \sigma_u^2 + \beta^2\text{var}(s) = \sigma_u^2 + \beta^2\Sigma_0 \\
E[y] &= E[x + u] = E[x] = E[\alpha + \beta s] = \alpha + \beta p_0.
\end{aligned}$$

Using these things we get that $E[v|y]$ is given by

$$E[v|y] = p_0 + \frac{\beta\Sigma_0}{\sigma_u^2 + \beta^2\Sigma_0}(y - \alpha - \beta p_0).$$

Setting this equal to $\lambda y + \mu$ we get equations for λ and μ

$$\lambda = \frac{\beta\Sigma_0}{\sigma_u^2 + \beta^2\Sigma_0} \tag{44}$$

$$\mu = p_0 - \frac{\beta\Sigma_0}{\sigma_u^2 + \beta^2\Sigma_0}(\alpha + \beta p_0). \tag{45}$$

Solving Equation 42, 43, 44, and 45 in `chapter_7_algebra.nb` for α, β, λ, and μ we get (when we take the root such that $\lambda > 0$)

$$\begin{aligned}
\alpha &= -\frac{p_0\sigma_u}{\sqrt{\Sigma_0 + 2\sigma_\epsilon^2}} \\
\beta &= \frac{\sigma_u}{\sqrt{\Sigma_0 + 2\sigma_\epsilon^2}} \\
\mu &= p_0 \\
\lambda &= \frac{\Sigma_0}{2\sigma_u\sqrt{\Sigma_0 + 2\sigma_\epsilon^2}}.
\end{aligned}$$

Exercise 7.3 (a piggy backing broker)

From the problem statement, the total demand ordered would be $x + \gamma x = (1+\gamma)x$, so the observed demand by the market market is $y = u + (1+\gamma)x$. The informed trader makes a profit given by

$$\pi = (v - p)x = (v - \lambda y - \mu)x = (v - \lambda(1+\gamma)x - \lambda u - \mu)x\,.$$

The expected profit, given that the informed trader knows the final price v, is

$$E[\pi|v] = (v - \lambda(1+\gamma)x - \mu)x\,.$$

This is the same objective function we have maximized before but now with $\lambda \to (1+\gamma)\lambda$. Thus the optimal x to order is

$$x = \frac{v-\mu}{2\lambda(1+\gamma)}\,.$$

Setting this equal to $\alpha + \beta v$ we get

$$\alpha = -\frac{\mu}{2\lambda(1+\gamma)} \tag{46}$$

$$\beta = \frac{1}{2\lambda(1+\gamma)}\,. \tag{47}$$

The market maker needs to compute $E[v|y] = E[v] + \frac{\sigma_{vy}}{\sigma_y^2}(y - E[y])$. As y can be expressed as

$$y = u + (1+\gamma)x = u + (1+\gamma)(\alpha + \beta v)\,,$$

the pieces we need to evaluate $E[v|y]$ are given by

$$\begin{aligned}
\sigma_{vy} &= \text{cov}(v, y) = \text{cov}(v, u + (1+\gamma)(\alpha+\beta v)) = (1+\gamma)\beta\Sigma_0 \\
\sigma_y^2 &= \text{var}(u + (1+\gamma)(\alpha+\beta v)) = \sigma_u^2 + (1+\gamma)^2\beta^2\Sigma_0 \\
E[y] &= (1+\gamma)(\alpha + \beta p_0)\,.
\end{aligned}$$

Thus we get for $E[v|y]$ we get

$$E[v|y] = p_0 + \left(\frac{(1+\gamma)\beta\Sigma_0}{\sigma_u^2 + (1+\gamma)^2\beta^2\Sigma_0}\right)(y - (1+\gamma)(\alpha+\beta p_0))\,.$$

Setting this expression equal to $\mu + \lambda y$ we get for λ and μ the following

$$\begin{aligned}
\lambda &= \frac{(1+\gamma)\beta\Sigma_0}{\sigma_u^2 + (1+\gamma)^2\beta^2\Sigma_0} \\
\mu &= p_0 - \frac{(1+\gamma)^2\beta\Sigma_0(\alpha+\beta p_0)}{\sigma_u^2 + (1+\gamma)^2\beta^2\Sigma_0}\,.
\end{aligned}$$

Using these two equations with α and β given by Equations 46 and 47 in `chapter_7_algebra.nb` we solve for the four values α, β, μ, and λ under the condition that $\lambda > 0$ to get

$$\begin{aligned}
\alpha &= -\frac{p_0\sigma_u}{(1+\gamma)\sqrt{\Sigma_0}} \\
\beta &= \frac{\sigma_u}{(1+\gamma)\sqrt{\Sigma_0}} \\
\mu &= p_0 \\
\lambda &= \frac{\sqrt{\Sigma_0}}{2\sigma_u}\,.
\end{aligned}$$

Chapter 8 (A Generalized Roll Model)

Notes on the text

The structural model

For the model

$$\begin{aligned} m_t &= m_{t-1} + w_t \\ w_t &= \lambda q_t + u_t \,. \end{aligned} \tag{48}$$

When we write it as

$$m_t = m_{t-1} + \lambda q_t + u_t \,, \tag{49}$$

we can more easily see the two contributions that affect the fair price m_t. The public information comes from u_t and the information from the informed traders come from the term λq_t. If we assume that a buy trade takes place (lift ones offer) then the trace price is

$$p_t = m_t + c = m_{t-1} + w_t + c = m_{t-1} + \lambda + c + u_t \,.$$

If a trade takes place on the bid (hit the bid)

$$p_t = m_t - c = m_{t-1} + w_t - c = m_{t-1} - \lambda - c + u_t \,.$$

Subtracting these two expressions gives that the bid-ask spread is given by

$$2(c + \lambda) \,.$$

This spread has two components now the $2c$ and 2λ.

Notes on the statistical representation of the generalized Roll model

When the trade price p_t is written as $p_t = m_t + cq_t$ and using the model given by Equation 49 for m_t we have that the *change* in the trade price is given by

$$\begin{aligned} \Delta p_t &= p_t - p_{t-1} = m_t + cq_t - m_{t-1} - cq_{t-1} \\ &= m_{t-1} + \lambda q_t + u_t + cq_t - m_{t-1} - cq_{t-1} = c(q_t - q_{t-1}) + \lambda q_t + u_t \,. \end{aligned} \tag{50}$$

Now $E[\Delta p_t] = 0$ since everything on the right-hand-side of the expression for Δp_t has zero mean. Squaring Δp_t we find

$$\begin{aligned} \Delta p_t^2 &= c^2(q_t - q_{t-1})^2 + \lambda c(q_t^2 - q_t q_{t-1}) + cu_t(q_t - q_{t-1}) \\ &+ \lambda c(q_t^2 - q_t q_{t-1}) + \lambda^2 q_t^2 + \lambda q_t u_t + cu_t(q_t - q_{t-1}) + \lambda q_t u_t + u_t^2 \,. \end{aligned}$$

To evaluate $E[\Delta p_t^2]$ using the above expression we will need the facts that

$$E[q_t^2] = 1^2 P\{q_t = +1\} + (-1)^2 P\{q_t = -1\} = 1 \,,$$

and facts like $E[q_t q_{t-1}] = E[q_t u_t] = 0$ etc. Then taking the expectation of Δp_t^2 then gives

$$
\begin{aligned}
E[\Delta p_t^2] &= c^2(2) + \lambda c + \lambda c + \lambda^2 + \sigma_u^2 \\
&= c^2 + (c+\lambda)^2 + \sigma_u^2 \,,
\end{aligned}
\tag{51}
$$

when we simplify. The above is the definition of γ_0. Now to evaluate $\gamma_1 \equiv E[\Delta p_t \Delta p_{t-1}]$ we first compute

$$
\begin{aligned}
\Delta p_t \Delta p_{t-1} &= (c(q_t - q_{t-1}) + \lambda q_t + u_t)(c(q_{t-1} - q_{t-2}) + \lambda q_{t-1} + u_{t-1}) \\
&= c^2(q_t - q_{t-1})(q_{t-1} - q_{t-2}) + \lambda c q_{t-1}(q_t - q_{t-1}) + c u_{t-1}(q_t - q_{t-1}) \\
&+ \lambda c q_t (q_{t-1} - q_{t-2}) + \lambda^2 q_t q_{t-1} + \lambda q_t u_{t-1} \\
&+ c u_t (q_{t-1} - q_{t-2}) + \lambda u_t q_{t-1} + u_t u_{t-1} \,.
\end{aligned}
$$

Taking the expectation of this we get

$$
E[\Delta p_t \Delta p_{t-1}] = c^2(-1) + \lambda c(-1) = -c(c+\lambda) \,. \tag{52}
$$

All later autocorrelations are zero. From the Wold theorem the model for Δp_t must be represented as a MA(1) model. In other words it can be written in the form

$$
\Delta p_t = \varepsilon_t + \theta \varepsilon_{t-1} \,. \tag{53}
$$

If we wish to apply the generalized Roll model to the changes in trade prices we could empirically measure Δp_t and then fit a MA(1) model to these prices changes. Fitting this MA(1) model will determine θ and σ_ε^2 empirically. Thus these two parameters θ and σ_ε^2 will be used below when needed.

Using the expression for w_t from Equation 48 we have

$$
\text{Var}(w_t) = \lambda^2(1) + \sigma_u^2 \,.
$$

Using Equation 51 for γ_0 and Equation 52 for γ_1 we can write

$$
\begin{aligned}
\gamma_0 &= 2c^2 + +2c\lambda + \lambda^2 + \sigma_u^2 \\
\gamma_1 &= -c^2 - c\lambda \,.
\end{aligned}
$$

The two of these equations show that $\gamma_0 + 2\gamma_1 = \lambda^2 + \sigma_u^2 = \text{Var}(w_t)$.

Note on Forecasting and Filtering

In the generalized Roll model the forecast prices are

$$
f_t = p_t + \theta \varepsilon_t \,. \tag{54}
$$

Here f_t is the forecast/filtered/smoothed trade price based on statistics of changes in trade prices or Δp_t. That is we assume a MA(1) model for Δp_t of the form $\Delta p_t = \varepsilon_t + \theta \varepsilon_{t-1}$ and from historic data of trade prices estimate θ and σ_ε^2. We will show that the expectation of m_t has a similar relationship. This is a useful representation because f_t computed in this way is a better estimate of m_t than the last trade price p_t is. We can use the form of the model for Δp_t to estimate ε_t as time progress, using the model written as $\varepsilon_t = \Delta p_t - \theta\varepsilon_{t-1}$. That is given the measurable sequence of trade price changes

$$\Delta p_1 , \quad \Delta p_2 , \quad \Delta p_3 , \quad \Delta p_4 \quad \cdots$$

We start our measurements of ε_t by assuming that $\varepsilon_0 = 0$ and then form the estimates of ε_t from

$$\begin{aligned}
\hat{\varepsilon}_1 &= \Delta p_1 - \theta(0) = \Delta p_1 \\
\hat{\varepsilon}_2 &= \Delta p_2 - \theta\hat{\varepsilon}_1 \\
\hat{\varepsilon}_3 &= \Delta p_3 - \theta\hat{\varepsilon}_2 \\
&\vdots
\end{aligned}$$

Thus at each instant of time t we observe the most recent trade price p_t, compute the change in trade price $\Delta p_t = p_t - p_{t-1}$, and then compute $\hat{\varepsilon}_t = \Delta p_t - \theta\hat{\varepsilon}_{t-1}$. Using this estimate we can use Equation 54 to compute the filtered estimate of fair. In any application where one needs accurate estimates of fair prices we could use the filtered estimate f_t. In the next section we will discuss how well f_t estimates m_t by computing estimates to $\text{Var}(p_t - m_t)$. This later estimate can be used any place an the uncertainty in ones fair price is needed.

Using the expression for Δp_t given via Equation 50 or

$$\Delta p_t = c(q_t - q_{t-1}) + \lambda q_t + u_t ,$$

and Wold's theorem since $\gamma_j = 0$ for $j \geq 2$ we know that Δp_t must be expressible as a MA(1) model. That is it has a representation given by $\Delta p_t = \varepsilon_t + \theta\varepsilon_{t-1}$ with some values (numbers) for the constants $(\theta, \sigma_\varepsilon^2)$. Thus from Equation 50 and this moving average representation we have

$$\varepsilon_t + \theta\varepsilon_{t-1} = (c + \lambda)q_t - cq_{t-1} + u_t . \tag{55}$$

or solving for ε_t we get

$$\varepsilon_t = (c + \lambda)q_t - cq_{t-1} + u_t - \theta\varepsilon_{t-1} .$$

Using this we see that $\text{Cov}(q_t, \varepsilon_t) = c + \lambda$ and $\text{Cov}(q_t, \varepsilon_{t-1}) = 0$ for all $k \geq 1$. With these expectations $E^*[q_t|\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots]$ can be calculated via our conditional expectation equation 34. We have for the expectation of q_t we find

$$\begin{aligned}
E^*[q_t|\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots] &= E^*[q_t|\varepsilon_t] \\
&= E[q_t] + \frac{\text{Cov}(q_t, \varepsilon_t)}{\sigma_\varepsilon^2}(\varepsilon_t - E[\varepsilon_t]) \\
&= 0 + \frac{(c + \lambda)}{\sigma_\varepsilon^2}(\varepsilon_t - 0) = \frac{(c + \lambda)}{\sigma_\varepsilon^2}\varepsilon_t .
\end{aligned}$$

Using this we can compute the filtered price f_t as

$$f_t = E^*[m_t|p_t, p_{t-1}, p_{t-2}, \ldots] = p_t - cE[q_t|\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots] = p_t - \frac{c(c + \lambda)}{\sigma_\varepsilon^2}\varepsilon_t .$$

fair price given all trades and the roll model general structure is given by Note that from Equation 52 we have

$$\gamma_1 = E(\Delta p_t, \Delta p_{t-1}) = -c(c+\lambda)\,,$$

while from the Wold representation of Δp_t i.e. $\Delta p_t = \varepsilon_t + \theta\varepsilon_{t-1}$ we can write the expectation in the definition of γ_1 as

$$E(\Delta p_t, \Delta p_{t-1}) = E(\varepsilon_t + \theta\varepsilon_{t-1}, \varepsilon_{t-1} + \theta\varepsilon_{t-2}) = \theta E[\varepsilon_{t-1}^2] = \theta\sigma_\varepsilon^2\,.$$

Solve for θ in these two expressions for $E(\Delta p_t, \Delta p_{t-1})$ we get

$$\theta = -\frac{c(c+\lambda)}{\sigma_\varepsilon^2}\,.$$

so

$$f_t = E^*[m_t|p_t, p_{t-1}, p_{t-2}, \ldots] = p_t + \theta\varepsilon_t\,. \tag{56}$$

is the expected fair price (filtered price) after observing the trade p_t. See notes earlier about how one would use these MA(1) representations to compute a better estimate of m_t.

Notes on the pricing error

In this section we want to study how well f_t estimate m_t in that we will consider the expression $\text{Var}(p_t - m_t)$. We can derive an expression for $\text{Var}(p_t - m_t)$ as

$$\text{Var}(p_t - m_t) = \text{Var}(m_t + cq_t - m_t) = \text{Var}(cq_t) = c^2\,.$$

The book states that we cannot compute c from the given time series data Δp_t. In that case we can try to compute a lower bound on this variance. We do this by writing

$$\text{Var}(p_t - m_t) = \text{Var}(p_t - f_t + f_t - m_t) = \text{Var}(p_t - f_t) + \text{Var}(f_t - m_t)\,.$$

We can use Equation 54 to compute the first term. To get a lower bound we will compute $\text{Var}(m_t - f_t)$ under a smaller amount of randomness i.e. we will take $u_t = 0$. This means that all information that changes our efficient fair price must be due to the trade λq_t information. In the original Roll model there is no trade information into m_t and we have $\lambda = 0$ so $m_t = m_{t-1} + u_t$ only where as here we are considering the case with no u_t and only trade information influencing the value of m_t. Using Equation 55 with $u_t = 0$ gives

$$\varepsilon_t + \theta\varepsilon_{t-1} = (c+\lambda)q_t - cq_{t-1}\,.$$

If we take $\varepsilon_t = (c+\lambda)q_t$ (by equating the expressions at time t) the remaining parts of the expression would need to be $\theta\varepsilon_{t-1} = -cq_{t-1}$. Incrementing t in this expression by one and using the previous relationship

$$\theta\varepsilon_t = -cq_t \quad \Rightarrow \quad \theta(c+\lambda)q_t = -cq_t\,.$$

Then to be consistent this means that

$$\theta = -\frac{c}{c+\lambda}$$

The filtered estimate of m_t is then

$$f_t = p_t + \theta\varepsilon_t = p_t - \frac{c}{c+\lambda}(c+\lambda)q_t = p_t - cq_t = m_t \,.$$

This means that when $\sigma_u^2 = 0$ the filtered estimate is exactly the same as the fair price so that $\text{Var}(m_t - f_t) = 0$. Using this we have

$$\begin{aligned}\text{Var}(p_t - m_t) &= \text{Var}(p_t - f_t) + \text{Var}(m_t - f_t) \\ &\geq \text{Var}(p_t - f_t) = \theta^2\sigma_\varepsilon^2 \equiv \underline{\sigma}_s^2 \,,\end{aligned}$$

and we have our *lower* bound on $\text{Var}(p_t - m_t)$ and have defined the expression $\underline{\sigma}_s^2$. Recall that θ and σ_ε^2 were fit to the Δp_t time series i.e. from the autocovariance of Δp_t

$$\begin{aligned}\gamma_0 &= E(\Delta p_t^2) = c^2 + (c+\lambda)^2 + \sigma_u^2 = (1+\theta^2)\sigma_\varepsilon^2 \\ \gamma_1 &= E(\Delta p_t \Delta p_{t-1}) = -c(c+\lambda) = \theta\sigma_\varepsilon^2 \,.\end{aligned}$$

Using the MA(1) representation given via Equations 15 and 16. Solving for the above for θ and σ_ε^2 we get Equations 17 (with the minus sign) and 18 (with the positive sign). When we multiply these expressions to compute $\underline{\sigma}_s^2 = \theta^2\sigma_\varepsilon^2$ we get

$$\theta^2\sigma_\varepsilon^2 = \frac{1}{2}(\gamma_0 - \sqrt{\gamma_0^2 - 4\gamma_1^2}) \,,$$

when we put in the expressions for γ_0 and γ_1 in terms of c and λ in the generalized Roll model we get

$$\underline{\sigma}_s^2 = \frac{1}{2}\left[c^2 + (c+\lambda)^2 + \sigma_u^2 - \sqrt{(\lambda^2 + \sigma_u^2)[(2c+\lambda)^2 + \sigma_u^2]}\right] \,, \tag{57}$$

the same expression as in the book. This is done in the Mathematica file `generalized_roll_model.nb`.

Notes on general univariate random-walk decompositions

From the representation of the filtered price f_t given in the book

$$f_t = p_t + \left(\sum_{j=0}^{\infty}\theta_{j+1}\right)\varepsilon_t + \left(\sum_{j=0}^{\infty}\theta_{j+2}\right)\varepsilon_{t-1} + \left(\sum_{j=0}^{\infty}\theta_{j+3}\right)\varepsilon_{t-2} + \cdots \,, \tag{58}$$

by decrementing the time index in f_t by one unit we get

$$f_{t-1} = p_{t-1} + \left(\sum_{j=0}^{\infty}\theta_{j+1}\right)\varepsilon_{t-1} + \left(\sum_{j=0}^{\infty}\theta_{j+2}\right)\varepsilon_{t-2} + \left(\sum_{j=0}^{\infty}\theta_{j+3}\right)\varepsilon_{t-3} + \cdots \,.$$

These two expressions can be subtracted when we line up terms containing ε that have the same time index

$$\begin{aligned}\Delta f_t &= \Delta p_t \\ &+ \left(\sum_{j=0}^{\infty}\theta_{j+1}\right)\varepsilon_t + \left(\sum_{j=0}^{\infty}\theta_{j+2} - \sum_{j=0}^{\infty}\theta_{j+1}\right)\varepsilon_{t-1} + \left(\sum_{j=0}^{\infty}\theta_{j+3} - \sum_{j=0}^{\infty}\theta_{j+2}\right)\varepsilon_{t-2} + \cdots \\ &= \Delta p_t + \left(\sum_{j=0}^{\infty}\theta_{j+1}\right)\varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \theta_3\varepsilon_{t-3} - \cdots \,.\end{aligned}$$

With the moving average representation for Δp_t of $\Delta p_t = \theta(L)\varepsilon_t = \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j}$ and the above, we can write Δf_t as

$$\Delta f_t = \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j} + \left(\sum_{j=0}^{\infty} \theta_{j+1} \right) \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_3 \varepsilon_{t-3} - \cdots \tag{59}$$

$$= \theta_0 \varepsilon_t + \left(\sum_{j=0}^{\infty} \theta_{j+1} \right) \varepsilon_t = \left(\sum_{j=0}^{\infty} \theta_j \right) \varepsilon_t = \theta(1)\varepsilon_t \,. \tag{60}$$

As the efficient price has dynamics given by $m_t = m_{t-1} + w_t$ we see that $\Delta m_t = w_t$. If we assert that $f_t \equiv m_t$ then from the above we would have

$$\Delta m_t = \Delta f_t = w_t = \theta(1)\varepsilon_t \,, \tag{61}$$

where the last equality used follows from Equation 60. Taking variance of both sides we get the structural requirement

$$\sigma_w^2 = \theta(1)^2 \sigma_\varepsilon^2 \,. \tag{62}$$

Using Equation 58 we get for the discrepancy between the trade prices and the efficient price m_t or $s_t = p_t - m_t$ the following

$$\begin{aligned} s_t &= p_t - m_t \\ &= f_t - \left(\sum_{j=0}^{\infty} \theta_{j+1} \right) \varepsilon_t - \left(\sum_{j=0}^{\infty} \theta_{j+2} \right) \varepsilon_{t-1} - \left(\sum_{j=0}^{\infty} \theta_{j+3} \right) \varepsilon_{t-2} + \cdots - m_t \,. \end{aligned}$$

If $f_t = m_t$ the first and last terms cancel and the above becomes

$$s_t = -\left(\sum_{j=0}^{\infty} \theta_{j+1} \right) \varepsilon_t - \left(\sum_{j=0}^{\infty} \theta_{j+2} \right) \varepsilon_{t-1} - \left(\sum_{j=0}^{\infty} \theta_{j+3} \right) \varepsilon_{t-2} + \cdots$$

$$= C_o \varepsilon_t + C_1 \varepsilon_{t-1} + C_2 \varepsilon_{t-2} + \cdots \,, \tag{63}$$

where C_i is given by

$$C_i \equiv -\sum_{j=i+1}^{\infty} \theta_j \,. \tag{64}$$

Using Equation 63 we have that

$$\sigma_s^2 = \sum_{i=0}^{\infty} C_i^2 \sigma_\varepsilon^2 \,. \tag{65}$$

Since we know that $\theta(1)\varepsilon_t = w_t$ we can write s_t which is expressed in terms of ε_t in terms of the equivalent w_t (up to a scaling by $\theta(1)^{-1}$) as

$$s_t = \sum_{i=0}^{\infty} C_i \varepsilon_{t-i} = \sum_{i=0}^{\infty} \frac{C_i}{\theta(1)} w_{t-i} \,.$$

Letting the coefficients of w_{t-i} be denoted as A_i so that

$$A_i = \frac{C_i}{\theta(1)} = -\frac{1}{\theta(1)} \sum_{j=i+1}^{\infty} \theta_j \,.$$

Exercise 8.1 (the Roll model with stale prices)

Here we assume that our trade prices are based on old efficient prices with the model $p_t = m_{t-1} + cq_t$. In this case we find for Δp_t the following

$$\Delta p_t = p_t - p_{t-1} = m_{t-1} + cq_t - m_{t-2} - cq_{t-1} = w_{t-1} + c(q_t - q_{t-1})\,.$$

From the given expression for Δp_t we compute

$$\begin{aligned}\gamma_0 = E[{\Delta p_t}^2] &= E[w_t^2] + 2E[cw_t(q_t - q_{t-1})] + c^2 E[(q_t - q_{t-1})^2] \\ &= \sigma_w^2 + 0 + c^2 E[q_t^2 - 2q_t q_{t-1} + q_{t-1}^2] \\ &= \sigma_w^2 + c^2(1) - 0 + c^2 = \sigma_w^2 + 2c^2\,.\end{aligned}$$

and

$$\begin{aligned}\gamma_1 = E[\Delta p_t \Delta p_{t-1}] &= E[(w_t + c(q_t - q_{t-1}))(w_{t-1} + c(q_{t-1} - q_{t-2}))] \\ &= 0 + c^2 E[(q_t - q_{t-1})(q_{t-1} - q_{t-2})] = c^2(-1) = -c^2\,.\end{aligned}$$

With $\gamma_k = 0$ for $k \geq 2$. Thus Δp_t would have a MA(2) representation.

Exercise 8.2 (lagged price adjustments)

When the trade prices satisfy $p_t = p_{t-1} + \alpha(m_t - p_{t-1})$ we find that

$$\begin{aligned}\Delta p_t = p_t - p_{t-1} &= p_{t-1} + \alpha(m_t - p_{t-1}) - p_{t-2} - \alpha(m_{t-1} - p_{t-2}) \\ &= \Delta p_{t-1} + \alpha(w_t - \Delta p_{t-1}) \\ &= (1-\alpha)\Delta p_{t-1} + \alpha w_t = (1-\alpha)L\Delta p_t + \alpha w_t\,.\end{aligned}$$

Thus the AR model for Δp_t is given by

$$(1 - (1-\alpha)L)\Delta p_t = \alpha w_t\,.$$

Solving for Δp_t in terms of its MA representation give $\Delta p_t = \phi(L)^{-1}\varepsilon_t$. We can then use Equation 61 as $w_t = \theta(1)\varepsilon_t$. Given what $\theta(1)$ is in this case i.e. $\theta(1) = \phi(1)^{-1}$ we would get $w_t = \phi(1)^{-1}\varepsilon_t$. Taking the variance of both sides of this expression gives

$$\sigma_w^2 = \phi(1)^{-2}\sigma_\varepsilon^2\,,$$

as expected.

Exercise 8.3 (variance calculations)

The given model for Δp_t is a MA(2) model. Here the MA(2) operator $\theta(L)$ is given by $\theta(L) = 1 - 0.3L + 0.1L^2$.

Part (a): We want to evaluate σ_w which we can do via Equation 62 which in this case is

$$\sigma_w^2 = (1 - 0.3 + 0.1)^2(0.00001) = 6.4\,10^{-6}\,.$$

Thus $\sigma_w = 0.0025$.

Part (b): As there are

$$N = \frac{60(6)}{5} = 72\,,$$

five minute intervals during a day we have that the variance of a day grows to 72 times σ_w^2 thus the standard deviation of a day is given by

$$\sqrt{72(6.4\,10^{-6})} = 0.0215\,.$$

Part (c): Using Equation 64 we find for this problem that

$$\begin{aligned}
C_0 &= -\theta_1 - \theta_2 = -(-0.3) - (0.1) = 0.2 \\
C_1 &= -\theta_2 = -0.1 \\
C_i &= 0 \quad \text{for} \quad i \geq 2\,.
\end{aligned}$$

Thus using Equation 65 we find

$$\sigma_s^2 = (C_0^2 + C_1^2)\sigma_\varepsilon^2 = (0.2^2 + 0.1^2)0.00001 = 5.0\,10^{-7}\,,$$

so $\sigma_s = 0.0007$ is the lower bound.

Notes on the Identification in Random-Walk Decompositions

From the form of $s_t = A(L)w_t + B(L)\eta_t$ we can compute

$$\Delta p_t = (1 + (1 - L)A(L))w_t + (1 - L)B(L)\eta_t\,. \tag{66}$$

Note the are two sources of noise in the above representation of Δp_t one from w_t and one from η_t. If we can write Δp_t in a MA representation as $\Delta p_t = \theta(L)\varepsilon_t$ with only one source of noise ε_t. We would have an autocovariance generating function given by

$$g_{\Delta p}(z) = \theta(z^{-1})\theta(z)\sigma_\varepsilon^2\,. \tag{67}$$

Which from the representation in Equation 66

$$g_{\Delta p}(z) = [1 + (1 - z^{-1})A(z^{-1})][1 + (1 - z)A(z)]\sigma_w^2 + (1 - z^{-1})B(z^{-1})(1 - z)B(z)\sigma_\eta^2\,. \tag{68}$$

Equating these two representation and letting $z = 1$ gives $\theta(1)^2\sigma_\varepsilon^2 = \sigma_w^2$.

Chapter 9 (Multivariate Linear Microstructure Models)

Notes on the text

The structural for prices and trades

Starting with serial autocorrelation of trade directions $q_t = \pm 1$ as a MA(1) process

$$q_t = v_t + \beta v_{t-1} , \tag{69}$$

so that the efficient price has dynamics

$$m_t = m_{t-1} + w_t \quad \text{with noise innovation driven by} \quad w_t = u_t + \lambda v_t . \tag{70}$$

Then changes in the trade price $p_t = m_t + cq_t$ is given by

$$\begin{aligned}
\Delta p_t &= p_t - p_{t-1} = m_t + cq_t - (m_{t-1} + cq_{t-1}) \\
&= m_{t-1} + w_t + cq_t - m_{t-1} - cq_{t-1} \\
&= u_t + \lambda v_t + c(q_t - q_{t-1}) \\
&= u_t + \lambda v_t + c(v_t + \beta v_{t-1} - v_{t-1} - \beta v_{t-2}) .
\end{aligned} \tag{71}$$

Using this expression with the dynamics for q_t gives for the variables Δp_t and q_t the system

$$\begin{aligned}
\Delta p_t &= u_t + (\lambda + c)v_t + c(\beta - 1)v_{t-1} - \beta c v_{t-2} \\
q_t &= v_t + \beta v_{t-1} .
\end{aligned}$$

These expressions show how Δp_t and q_t can be written in terms of current and lagged innovations u_t and v_t. In vector form the above can be written as

$$\begin{bmatrix} \Delta p_t \\ q_t \end{bmatrix} = \begin{bmatrix} 1 & \lambda + c \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_t \\ v_t \end{bmatrix} + \begin{bmatrix} 0 & c(\beta - 1) \\ 0 & \beta \end{bmatrix} \begin{bmatrix} u_{t-1} \\ v_{t-1} \end{bmatrix} + \begin{bmatrix} 0 & -\beta c \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_{t-2} \\ v_{t-2} \end{bmatrix} . \tag{72}$$

Let the vector ε_t be defined as $\varepsilon_t \equiv \begin{bmatrix} 1 & \lambda + c \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_t \\ v_t \end{bmatrix}$ so that

$$\begin{bmatrix} u_t \\ v_t \end{bmatrix} = \begin{bmatrix} 1 & \lambda + c \\ 0 & 1 \end{bmatrix}^{-1} \varepsilon_t = \begin{bmatrix} 1 & -(\lambda + c) \\ 0 & 1 \end{bmatrix} \varepsilon_t .$$

Then to write our model for $\begin{bmatrix} \Delta p_t & q_t \end{bmatrix}$ in terms of ε_t rather than $\begin{bmatrix} u_t & v_t \end{bmatrix}$ we need to compute the matrix products

$$\begin{aligned}
\begin{bmatrix} 0 & c(\beta - 1) \\ 0 & \beta \end{bmatrix} \begin{bmatrix} 1 & -\lambda - c \\ 0 & 1 \end{bmatrix} &= \begin{bmatrix} 0 & c(\beta - 1) \\ 0 & \beta \end{bmatrix} \equiv \theta_1 \\
\begin{bmatrix} 0 & -\beta c \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -\lambda - c \\ 0 & 1 \end{bmatrix} &= \begin{bmatrix} 0 & -\beta c \\ 0 & 0 \end{bmatrix} \equiv \theta_2 .
\end{aligned}$$

In the above I have defined the two matrices θ_1 and θ_2 that are found in the VMA(2) representation of $\left[\begin{array}{cc} \Delta p_t & q_t \end{array}\right]$ as

$$\left[\begin{array}{c} \Delta p_t \\ q_t \end{array}\right] = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \,.$$

Using the above transformational definition of $\varepsilon_t = B \left[\begin{array}{c} u_t \\ v_t \end{array}\right]$ we have

$$\begin{aligned}
\Omega = \text{Var}(\varepsilon_t) = \text{Var}\left(B \left[\begin{array}{c} u_t \\ v_t \end{array}\right] \right) &= B \text{Var}\left(\left[\begin{array}{c} u_t \\ v_t \end{array}\right] \right) B' \\
&= \left[\begin{array}{cc} 1 & c+\lambda \\ 0 & 1 \end{array}\right] \left[\begin{array}{cc} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{array}\right] \left[\begin{array}{cc} 1 & 0 \\ c+\lambda & 1 \end{array}\right] = \left[\begin{array}{cc} 1 & c+\lambda \\ 0 & 1 \end{array}\right] \left[\begin{array}{cc} \sigma_u^2 & 0 \\ \sigma_v^2 (c+\lambda) & \sigma_v^2 \end{array}\right] \\
&= \left[\begin{array}{cc} \sigma_u^2 + \sigma_v^2 (c+\lambda)^2 & \sigma_v^2 (c+\lambda) \\ \sigma_v^2 (c+\lambda) & \sigma_v^2 \end{array}\right] .
\end{aligned} \tag{73}$$

To compute the vector autoregressive (VAR) representation or $\phi(L)$ in terms of the vector moving average representation (VMA) $\theta(L)$ we need invert the polynomial $\theta(L) = I + \theta_1 L + \theta_2 L^2$ above as

$$\theta(L)^{-1} = (I + \theta_1 L + \theta_2 L^2)^{-1} = I - \phi_1 L - \phi_2 L^2 - \phi_3 L^3 - \phi_4 L^4 - \cdots \,.$$

If we multiply by $I + \theta_1 L + \theta_2 L^2$ on the right-hand-side we get the polynomial expression

$$I = (I - \phi_1 L - \phi_2 L^2 - \phi_3 L^3 - \phi_4 L^4 - \cdots)(I + \theta_1 L + \theta_2 L^2) \,.$$

Expanding the right-hand-side of this expression and carefully aligning each term

$$\begin{array}{cccccccc}
I = & I & -\phi_1 L & -\phi_2 L^2 & -\phi_3 L^3 & -\phi_4 L^4 & - \cdots \\
+ & & \theta_1 L & -\phi_1 \theta_1 L^2 & -\phi_2 \theta_1 L^3 & -\phi_3 \theta_1 L^4 & - \cdots \\
+ & & & +\theta_2 L^2 & -\phi_1 \theta_2 L^3 & -\phi_2 \theta_2 L^4 & - \cdots
\end{array}$$

Equating the coefficients of the powers of L^1 between each side we get

$$0 = -\phi_1 + \theta_1 \quad \Rightarrow \quad \phi_1 = \theta_1 = \left[\begin{array}{cc} 0 & c(\beta - 1) \\ 0 & \beta \end{array}\right] .$$

For the coefficients of L^2 we get $-\phi_2 - \phi_1 \theta_1 + \theta_2 = 0$ so

$$\begin{aligned}
\phi_2 &= -\phi_1 \theta_1 + \theta_2 \\
&= -\left[\begin{array}{cc} 0 & c(\beta - 1) \\ 0 & \beta \end{array}\right] \left[\begin{array}{cc} 0 & c(\beta - 1) \\ 0 & \beta \end{array}\right] + \left[\begin{array}{cc} 0 & -\beta c \\ 0 & 0 \end{array}\right] \\
&= -\left[\begin{array}{cc} 0 & c\beta(\beta - 1) \\ 0 & \beta^2 \end{array}\right] + \left[\begin{array}{cc} 0 & -\beta c \\ 0 & 0 \end{array}\right] = \left[\begin{array}{cc} 0 & -c\beta^2 \\ 0 & -\beta^2 \end{array}\right] .
\end{aligned}$$

For the coefficients of L^3 we get $-\phi_3 - \phi_2 \theta_1 - \phi_1 \theta_2 = 0$ so

$$\begin{aligned}
\phi_3 &= -\phi_2 \theta_1 - \phi_1 \theta_2 \\
&= -\left[\begin{array}{cc} 0 & -c\beta^2 \\ 0 & -\beta^2 \end{array}\right] \left[\begin{array}{cc} 0 & c(\beta - 1) \\ 0 & \beta \end{array}\right] - \left[\begin{array}{cc} 0 & c(\beta - 1) \\ 0 & \beta \end{array}\right] \left[\begin{array}{cc} 0 & -\beta c \\ 0 & 0 \end{array}\right] \\
&= -\left[\begin{array}{cc} 0 & -c\beta^3 \\ 0 & -\beta^3 \end{array}\right] - \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array}\right] = \left[\begin{array}{cc} 0 & c\beta^3 \\ 0 & \beta^3 \end{array}\right] .
\end{aligned}$$

In general the equation for the coefficients of L^k is $-\phi_k - \phi_{k-1} \theta_1 - \phi_{k-2} \theta_2 = 0$ or

$$\phi_k = -\phi_{k-1} \theta_1 - \phi_{k-2} \phi_2 \,.$$

Notes on forecasts and impulse response functions

To compute forecasts of the vector y_t we use

$$\begin{aligned}E^*[y_{t+k}|y_t, y_{t-1}, \cdots] &= E^*[y_{t+k}|\varepsilon_t, \varepsilon_{t-1}, \cdots] \\ &= E^*[\varepsilon_{t+k} + \theta_1\varepsilon_{t+k-1} + \theta_2\varepsilon_{t+k-2} + \cdots + \theta_{k-1}\varepsilon_{t+1} + \theta_k\varepsilon_t + \theta_{k+1}\varepsilon_{t-1} + \cdots|\varepsilon_t, \varepsilon_{t-1}, \cdots] \\ &= \theta_k\varepsilon_t + \theta_{k+1}\varepsilon_{t-1} + \theta_{k+2}\varepsilon_{t-2} + \cdots . \qquad (74)\end{aligned}$$

Where we have used the fact that

$$E^*[\varepsilon_{t+l}|\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \cdots] = 0 \quad \text{forall} \quad l \geq 1 .$$

The expected fair price is computed using

$$\begin{aligned}f_t &= \lim_{k\to\infty} E^*[p_{t+k}|p_t, p_{t-1}] \\ &= \lim_{k\to\infty} E^*[p_t + (p_{t+1} - p_t) + (p_{t+2} - p_{t+1}) + \cdots + (p_{t+k-1} - p_{t+k-2}) + (p_{t+k} - p_{t+k-1})|p_t, p_{t-1}] \\ &= \lim_{k\to\infty} E^*[p_t + \Delta p_{t+1} + \Delta p_{t+2} + \cdots + \Delta p_{t+k-1} + \Delta p_{t+k}|p_t, p_{t-1}] \\ &= \lim_{k\to\infty} E^*\left[p_t + \sum_{j=1}^{k} \Delta p_{t+j} \middle| p_t, p_{t-1}\right] = p_t + \lim_{k\to\infty} E^*\left[\sum_{j=1}^{k} \Delta p_{t+j} \middle| p_t, p_{t-1}\right] \\ &= p_t + E^*\left[\sum_{j=1}^{\infty} \Delta p_{t+j} \middle| p_t, p_{t-1}\right] = p_t + \sum_{j=1}^{\infty} E^*[\Delta p_{t+j}|p_t, p_{t-1}] .\end{aligned}$$

The structural VMA model of the stacked vector y_t where $y_t = \begin{bmatrix} \Delta p_t \\ x_t \end{bmatrix}$ where Δp_t is the first component is given by Eq. 9.2 from the book

$$y_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3} + \cdots = (I + \theta_1 L + \theta_2 L^2 + \theta_3 L^3 + \cdots)\varepsilon_t . \qquad (75)$$

When we increment t in Equation 75 by k to get y_{t+k} and take the first component i.e. the first row to get Δp_{t+k} then using Equation 74 we find

$$\begin{aligned}f_t &= p_t + E^*[\Delta p_{t+1}|p_t, p_{t-1}] + E^*[\Delta p_{t+2}|p_t, p_{t-1}] + E^*[\Delta p_{t+3}|p_t, p_{t-1}] + \cdots \\ &= p_t \\ &+ [\theta_1]_1\varepsilon_t + [\theta_2]_1\varepsilon_{t-1} + [\theta_3]_1\varepsilon_{t-2} + \cdots \\ &+ [\theta_2]_1\varepsilon_t + [\theta_3]_1\varepsilon_{t-1} + [\theta_4]_1\varepsilon_{t-2} + \cdots \\ &+ [\theta_3]_1\varepsilon_t + [\theta_4]_1\varepsilon_{t-1} + [\theta_5]_1\varepsilon_{t-2} + \cdots \\ &= p_t + \left(\sum_{j=0}^{\infty}[\theta_{j+1}]_1\right)\varepsilon_t + \left(\sum_{j=0}^{\infty}[\theta_{j+2}]_1\right)\varepsilon_{t-1} + \left(\sum_{j=0}^{\infty}[\theta_{j+3}]_1\right)\varepsilon_{t-2} + \cdots . \qquad (76)\end{aligned}$$

The notation $[\cdot]_1$ means to take the first component (row) of its argument.

The impulse response function $\psi_s(\varepsilon_0)$ is defined as

$$\psi_s(\varepsilon_0) = E^*[y_s|\varepsilon_0, \varepsilon_{-1} = \varepsilon_{-2} = \varepsilon_{-3} = \cdots = 0] , \qquad (77)$$

for $s \geq 0$. Thus we consider the expected response at s when our system is started with the initial value of ε_0. This is that only the first innovation is nonzero and all other innovations are zero. For the VMA model given by Equation 75 written for y_s we have

$$y_s = \varepsilon_s + \theta_1 \varepsilon_{s-1} + \theta_2 \varepsilon_{s-2} + \theta_3 \varepsilon_{s-3} + \cdots + \theta_{s-1} \varepsilon_1 + \theta_s \varepsilon_0 + \theta_{s+1} \varepsilon_{-1} + \theta_{s+2} \varepsilon_{-2} + \cdots$$

If we assume that $\varepsilon_{-1} = \varepsilon_{-2} = \varepsilon_{-3} = \cdots = 0$ then all terms after $\theta_s \varepsilon_0$ are zero and all terms after $\theta_s \varepsilon_0$ are not observed (and have expectation 0). Thus the impulse response function for a VMA model is

$$\psi_s(\varepsilon_0) = E^*[y_s | \varepsilon_0, \varepsilon_{-1} = \varepsilon_{-2} = \varepsilon_{-3} = \cdots = 0] = \theta_s \varepsilon_0 \, .$$

Notes on resolution of contemporaneous effects

Here we verify that the F given via

$$F = \begin{bmatrix} \sigma_1 & \sigma_{12}/\sigma_1 \\ 0 & \sqrt{\sigma_2^2 - \sigma_{12}^2/\sigma_1^2} \end{bmatrix} , \quad (78)$$

does in fact give the Cholesky factorization of the matrix $\begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$. We find the product of $F'F$ given by

$$\begin{aligned} F'F &= \begin{bmatrix} \sigma_1 & 0 \\ \sigma_{12}/\sigma_1 & \sqrt{\sigma_2^2 - \sigma_{12}^2/\sigma_1^2} \end{bmatrix} \begin{bmatrix} \sigma_1 & \sigma_{12}/\sigma_1 \\ 0 & \sqrt{\sigma_2^2 - \sigma_{12}^2/\sigma_1^2} \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_{12}^2/\sigma_1^2 + \sigma_2^2 - \sigma_{12}^2/\sigma_1^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} , \end{aligned}$$

which is what we wanted to show. In the two-dimensional case the Cholesky factorization provides a way to generate the vector $\mathbf{x}$ from the vector $\mathbf{v}$. For example, x_1 and x_2 are generated from z_1 and z_2 using the transformation

$$\begin{aligned} x_1 &= \sigma_1 z_1 \\ x_2 &= \frac{\sigma_{12}}{\sigma_1} z_1 + \sqrt{\sigma_2^2 - \frac{\sigma_{12}^2}{\sigma_1^2}} z_2 \, . \end{aligned}$$

Thus in the above factorization the randomness in z_1 first is used to generate x_1 and then that randomness is feed into the how the value of x_2 is computed. Thus information about x_1 is used in the computation of x_2 and information flows from x_1 to x_2. From this very simple argument, if we desire to study causal effects i.e. the variable x causes the variable y then the ordering of variables matters since the Cholesky factorization introduces an information flow. In the order of the variables developed thus far or $\begin{bmatrix} \Delta p_t & q_t \end{bmatrix}^T$ as Δp_t is the first variable, the Cholesky factorization states that the change in price is informative in determining the trade direction. We would expect information to flow in the other order. Thus we might reconsider the developments performed thus far in the other order or $\begin{bmatrix} q_t & \Delta p_t \end{bmatrix}$. This would affect the steps around Equation 72. In fact Equation 72 in this new variable ordering becomes

$$\begin{bmatrix} q_t \\ \Delta p_t \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & \lambda + c \end{bmatrix} \begin{bmatrix} u_t \\ v_t \end{bmatrix} + \begin{bmatrix} 0 & \beta \\ 0 & c(\beta - 1) \end{bmatrix} \begin{bmatrix} u_{t-1} \\ v_{t-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & -\beta c \end{bmatrix} \begin{bmatrix} u_{t-2} \\ v_{t-2} \end{bmatrix} .$$

Then with this order we would have

$$\varepsilon_t^* = \begin{bmatrix} \varepsilon_{q,t} \\ \varepsilon_{\Delta p,t} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & c + \lambda \end{bmatrix} \begin{bmatrix} u_t \\ v_t \end{bmatrix} = \begin{bmatrix} v_t \\ u_t + (c + \lambda) v_t \end{bmatrix} \tag{79}$$

With this definition of the disturbance we have a covariance matrix Ω^* given by

$$\begin{aligned} \Omega^* &= \begin{bmatrix} 0 & 1 \\ 1 & c + \lambda \end{bmatrix} \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & c + \lambda \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & c + \lambda \end{bmatrix} \begin{bmatrix} 0 & \sigma_u^2 \\ \sigma_v^2 & \sigma_v^2(c + \lambda) \end{bmatrix} \\ &= \begin{bmatrix} \sigma_v^2 & (c + \lambda)\sigma_v^2 \\ (c + \lambda)\sigma_v^2 & \sigma_u^2 + (c + \lambda)^2 \sigma_v^2 \end{bmatrix} . \end{aligned} \tag{80}$$

Then the Cholesky factorization of Ω^* or F^* via Equation 78 is given by

$$F = \begin{bmatrix} \sigma_v & (c + \lambda)\sigma_v^2 / \sigma_v \\ 0 & \sqrt{\sigma_u^2 + (c + \lambda)^2 \sigma_v^2 - (c + \lambda)^2 \sigma_v^4 / \sigma_v^2} \end{bmatrix} = \begin{bmatrix} \sigma_v & (c + \lambda)\sigma_v \\ 0 & \sigma_u \end{bmatrix} . \tag{81}$$

Warning: I had trouble deriving the matrix F that is a type of Cholesky factor of Ω (it is not upper triangular like a Cholesky factor should be). I'll present what I attempted here which did not give the same results as in the book. If anyone sees anything wrong with what I attempted here, please contact me.

To return to the starting order for the variable $\begin{bmatrix} \Delta p_t & q_t \end{bmatrix}^T$ means that we need to exchange the order of the elements in ε_t^* so

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{\Delta p,t} \\ \varepsilon_{q,t} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{q,t} \\ \varepsilon_{\Delta p,t} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \varepsilon_t^* .$$

This to me means that the variance of ε_t should be given by (when we use the Cholesky factorization of Ω^*)

$$\begin{aligned} \mathrm{Var}(\varepsilon_t) &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \mathrm{Var}(\varepsilon_t^*) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \Omega^* \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} F^{*\prime} F^* \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \left(F^* \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right)' \left(F^* \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right) \end{aligned}$$

Thus to set $\Omega = F'F$ we see from the above that we can take F given by

$$F = F^* \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \sigma_v & (c + \lambda)_v \\ 0 & \sigma_u \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} (c + \lambda)\sigma_v & \sigma_v \\ \sigma_u & 0 \end{bmatrix} .$$

The product of the above matrix F in the form $F'F$ does in fact equal Ω given via Equation 73 as it must but the matrix F computed above is not upper triangular as it should be to be a

Cholesky factor (as given by the definition) and it does not match the result for F given in the book. The F matrix that is given in the book (with the errata transpose from the web) is

$$F = \begin{bmatrix} \sigma_u & 0 \\ (c+\lambda)\sigma_v & \sigma_v \end{bmatrix} .$$

The product of the above matrix F in the form $F'F$ again does equal Ω given via Equation 73 as it must but this F is also not upper triangular either (it is lower triangular). At this point I gave up as not able to derive the result in the book. As always, if anyone sees an error in what have done please contact me.

Notes on a random walk variance

From the expression

$$\sigma_w^2 = [\theta(1)]_1 \Omega [\theta(1)]_1' \, , \tag{82}$$

note that w is a scalar so $[\theta(1)]_1$ is a row vector so $[\theta(1)]_1'$ is a column vector. For the model considered with vector unknown given by $\begin{bmatrix} \Delta p_t & q_t \end{bmatrix}^T$ we have

$$\theta(L) = I + \begin{bmatrix} 0 & c(\beta-1) \\ 0 & \beta \end{bmatrix} L + \begin{bmatrix} 0 & -c\beta \\ 0 & 0 \end{bmatrix} L^2 \quad \text{so} \quad \theta(1) = \begin{bmatrix} 1 & -c \\ 0 & \beta \end{bmatrix} ,$$

and we have $[\theta(1)]_1 = \begin{bmatrix} 1 & -c \end{bmatrix}$. Using this we can compute σ_w^2 using Equation 82 as

$$\begin{aligned} \sigma_w^2 &= \begin{bmatrix} 1 & -c \end{bmatrix} \begin{bmatrix} \sigma_u^2 + \sigma_v^2(c+\lambda)^2 & \sigma_v^2(c+\lambda) \\ \sigma_v^2(c+\lambda) & \sigma_v^2 \end{bmatrix} \begin{bmatrix} 1 \\ -c \end{bmatrix} \\ &= \begin{bmatrix} 1 & -c \end{bmatrix} \begin{bmatrix} \sigma_u^2 + (c+\lambda)^2\sigma_v^2 - (c^2 + c\lambda)\sigma_v^2 \\ (c+\lambda)\sigma_v^2 - c\sigma_v^2 \end{bmatrix} = \begin{bmatrix} 1 & -c \end{bmatrix} \begin{bmatrix} \sigma_u^2 + (c\lambda + \lambda^2)\sigma_v^2 \\ \lambda\sigma_v^2 \end{bmatrix} \\ &= \sigma_u^2 + (c\lambda + \lambda^2)\sigma_v^2 - c\lambda\sigma_v^2 = \sigma_u^2 + \lambda^2\sigma_v^2 \, . \end{aligned} \tag{83}$$

Notes on the pricing error

In this section we want to estimate the variance of the pricing error s_t where s_t is defined as $s_t = p_t - f_t$. From Equation 76 the formula for f_t we have that

$$s_t = p_t - f_t = -\left(\sum_{j=0}^{\infty} [\theta_{j+1}]_1 \right) \varepsilon_t - \left(\sum_{j=0}^{\infty} [\theta_{j+2}]_1 \right) \varepsilon_{t-1} - \left(\sum_{j=0}^{\infty} [\theta_{j+3}]_1 \right) \varepsilon_{t-2} + \cdots .$$

Since each of these terms is independent $E\varepsilon_t\varepsilon_{t-k}' = 0$ for $k \neq 0$ so we get

$$\sigma_s^2 = \text{Var}\left(\sum_{j=0}^{\infty} [\theta_{j+1}]_1 \varepsilon_t \right) + \text{Var}\left(\sum_{j=0}^{\infty} [\theta_{j+2}]_1 \varepsilon_{t-1} \right) + \text{Var}\left(\sum_{j=0}^{\infty} [\theta_{j+3}]_1 \varepsilon_{t-2} \right) + \cdots .$$

Lets evaluate one of the terms in the above sum, say the first term. Since the expression $\sum_{j=0}^{\infty}[\theta_{j+1}]_1$ is just a vector multiple of ε_t we find

$$\mathrm{Var}\left(\sum_{j=0}^{\infty}[\theta_{j+1}]_1\varepsilon_t\right) = \left(\sum_{j=0}^{\infty}[\theta_{j+1}]_1\right)\mathrm{Var}(\varepsilon_t)\left(\sum_{j=0}^{\infty}[\theta_{j+1}]_1\right)' = \left(\sum_{j=0}^{\infty}[\theta_{j+1}]_1\right)\Omega\left(\sum_{j=0}^{\infty}[\theta_{j+1}]_1\right)' .$$

Thus the expression for σ_s^2 becomes

$$\begin{aligned}
\sigma_s^2 &= \left(\sum_{j=0}^{\infty}[\theta_{j+1}]_1\right)\Omega\left(\sum_{j=0}^{\infty}[\theta_{j+1}]_1\right) + \left(\sum_{j=0}^{\infty}[\theta_{j+2}]_1\right)\Omega\left(\sum_{j=0}^{\infty}[\theta_{j+2}]_1\right)' + \left(\sum_{j=0}^{\infty}[\theta_{j+3}]_1\right)\Omega\left(\sum_{j=0}^{\infty}[\theta_{j+3}]_1\right)' + \cdots \\
&= \left(\sum_{j=1}^{\infty}[\theta_j]_1\right)\Omega\left(\sum_{j=1}^{\infty}[\theta_j]_1\right) + \left(\sum_{j=2}^{\infty}[\theta_j]_1\right)\Omega\left(\sum_{j=2}^{\infty}[\theta_j]_1\right)' + \left(\sum_{j=3}^{\infty}[\theta_j]_1\right)\Omega\left(\sum_{j=3}^{\infty}[\theta_j]_1\right)' + \cdots \\
&= \sum_{k=0}^{\infty}\left\{\left(\sum_{j=k+1}^{\infty}[\theta_j]_1\right)\Omega\left(\sum_{j=k+1}^{\infty}[\theta_j]_1\right)'\right\}
\end{aligned}$$

Define

$$C_k \equiv -\sum_{j=k+1}^{\infty}[\theta_j]_1\,, \tag{84}$$

(the negative sign but that does not matter in the evaluation of σ_s^2) and we get

$$\sigma_s^2 = \sum_{k=0}^{\infty}C_k\Omega C_k'\,. \tag{85}$$

In the example structural model considered in this chapter a *direct* calculation of σ_s^2 would give (if we assume $f_t = m_t$ (so as to compute a lower bound on σ_s^2)

$$s_t = p_t - f_t = m_t + cq_t - m_t = cq_t\,.$$

Now q_t in terms of moving average innovations gives

$$s_t = c(v_t + \beta v_{t-1})\,.$$

Thus

$$\sigma_s^2 = c^2(\sigma_v^2 + \beta^2\sigma_v^2) = c^2(1+\beta^2)\sigma_v^2\,.$$

We can check this result against the general multidimensional results derived above. We first need to compute C_0, C_1, C_2 etc. We find

$$\begin{aligned}
C_0 &= -\sum_{j=1}^{\infty}[\theta_j]_1 = -\sum_{j=1}^{2}[\theta_j]_1 = -\left(\left[\begin{array}{cc} 0 & c(\beta-1)\end{array}\right] + \left[\begin{array}{cc} 0 & -c\beta\end{array}\right]\right) \\
&= -\left[\begin{array}{cc} 0 & -c\end{array}\right] = \left[\begin{array}{cc} 0 & c\end{array}\right] \\
C_1 &= -\sum_{j=2}^{\infty}[\theta_j]_1 = -\left[\begin{array}{cc} 0 & -c\beta\end{array}\right] = \left[\begin{array}{cc} 0 & c\beta\end{array}\right] \\
C_j &= 0 \quad \text{for} \quad j \geq 2\,.
\end{aligned}$$

Thus using the above method we have

$$
\begin{aligned}
\underline{\sigma_s}^2 &= \left[\begin{array}{cc} 0 & c \end{array}\right] \Omega \left[\begin{array}{cc} 0 & c \end{array}\right]' + \left[\begin{array}{cc} 0 & c\beta \end{array}\right] \Omega \left[\begin{array}{cc} 0 & c\beta \end{array}\right]' \\
&= \left[\begin{array}{cc} 0 & c \end{array}\right] \left[\begin{array}{cc} \sigma_u^2 + \sigma_v^2(c+\lambda)^2 & \sigma_v^2(c+\lambda) \\ \sigma_v^2(c+\lambda) & \sigma_v^2 \end{array}\right] \left[\begin{array}{c} 0 \\ c \end{array}\right] \\
&+ \left[\begin{array}{cc} 0 & c\beta \end{array}\right] \left[\begin{array}{cc} \sigma_u^2 + \sigma_v^2(c+\lambda)^2 & \sigma_v^2(c+\lambda) \\ \sigma_v^2(c+\lambda) & \sigma_v^2 \end{array}\right] \left[\begin{array}{c} 0 \\ c\beta \end{array}\right] \\
&= \left[\begin{array}{cc} 0 & c \end{array}\right] \left[\begin{array}{c} c(c+\lambda)\sigma_v^2 \\ c\sigma_v^2 \end{array}\right] + \left[\begin{array}{cc} 0 & c\beta \end{array}\right] \left[\begin{array}{c} c\beta(c+\lambda)\sigma_v^2 \\ c\beta\sigma_v^2 \end{array}\right] \\
&= c^2\sigma_v^2 + c^2\beta^2\sigma_v^2 = c^2(1+\beta^2)\sigma_v^2 \,. \qquad (86)
\end{aligned}
$$

The same as earlier.

Problem 9.1 (the VMA model of GH)

The model we consider is given by

$$
\begin{aligned}
m_t &= m_{t-1} + w_t \\
w_t &= u_t + q_t(\lambda_0 + \lambda_1 V_t) \\
p_t &= m_t + q_t(c_0 + c_1 V_t) \,.
\end{aligned}
$$

Now the noise in this model comes from three sources. The to apply to the midpoint (aka the efficient price) come from u_t, q_t, and Q_t (which is assumed to be a normal random variable $Q_t \sim \mathcal{N}(0, \sigma_Q^2)$). Thus the disturbance vector ε_t is $\varepsilon_t = \left[\begin{array}{c} u_t \\ q_t \\ Q_t \end{array}\right]$.

Part (a): For this part we want to write $y_t \equiv \left[\begin{array}{c} \Delta p_t \\ q_t \\ Q_t \end{array}\right] = \theta(L)\varepsilon_t$. We compute Δp_t as

$$
\begin{aligned}
\Delta p_t = p_t - p_{t-1} &= m_t + q_t(c_0 + c_1 V_t) - m_{t-1} - q_{t-1}(c_0 + c_1 V_{t-1}) \\
&= w_t + c_0(q_t - q_{t-1}) + c_1(q_t V_t - q_{t-1} V_{t-1}) \\
&= w_t + c_0(q_t - q_{t-1}) + c_1(Q_t - Q_{t-1}) \quad \text{since } q_t V_t = Q_t \\
&= u_t + q_t(\lambda_0 + \lambda_1 V_t) + c_0(q_t - q_{t-1}) + c_1(Q_t - Q_{t-1}) \\
&= u_t + (\lambda_0 + c_0) q_t + (\lambda_1 + c_1) Q_t - c_0 q_{t-1} - c_1 Q_{t-1} \,.
\end{aligned}
$$

Thus our model becomes (in matrix form) using ε_t defined earlier

$$
\left[\begin{array}{c} \Delta p_t \\ q_t \\ Q_t \end{array}\right] = \left[\begin{array}{ccc} 1 & \lambda_0 + c_0 & \lambda_1 + c_1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c} u_t \\ q_t \\ Q_t \end{array}\right] + \left[\begin{array}{ccc} 0 & -c_0 & -c_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] \left[\begin{array}{c} u_{t-1} \\ q_{t-1} \\ Q_{t-1} \end{array}\right] ,
$$

Thus $\theta_0 = \left[\begin{array}{ccc} 1 & \lambda_0 + c_0 & \lambda_1 + c_1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]$ and $\theta_1 = \left[\begin{array}{ccc} 0 & -c_0 & -c_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$. We find for the requested covariance given by

$$\text{Cov}(q_t, Q_t) = E[(q_t - E[q_t])(Q_t - E[Q_t])'] = E[q_t Q_t] = E[|Q_t|] = \sqrt{\frac{2\sigma_Q^2}{\pi}},$$

using the identity given in the book. Next we compute

$$\Omega = \text{Var}(\varepsilon_t) = E[\varepsilon_t \varepsilon_t'] = E\left[\left[\begin{array}{c} u_t \\ q_t \\ Q_t \end{array}\right]\left[\begin{array}{ccc} u_t & q_t & Q_t \end{array}\right]\right]$$

$$= E\left[\left[\begin{array}{ccc} u_t^2 & u_t q_t & u_t Q_t \\ q_t u_t & q_t^2 & q_t Q_t \\ Q_t u_t & Q_t q_t & Q_t^2 \end{array}\right]\right] = \left[\begin{array}{ccc} \sigma_u^2 & 0 & 0 \\ 0 & 1 & \sqrt{\frac{2}{\pi}}\sigma_Q \\ 0 & \sqrt{\frac{2}{\pi}}\sigma_Q & \sigma_Q^2 \end{array}\right]$$

Part (b): We can compute σ_w^2 via Equation 86 as

$$w_t = u_t + q_t(\lambda_0 + \lambda_1 V_t) = u_t + \lambda_0 q_t + \lambda_1 Q_t .$$

As u_t, q_t, and Q_t are the independent innovations in this problem the variance of w_t is easy to compute (notice that w_t has a zero mean) so

$$\begin{aligned} \sigma_w^2 &= E[(u_t + \lambda_0 q_t + \lambda_1 Q_t)^2] \\ &= \sigma_u^2 + E[(\lambda_0 q_t + \lambda_1 Q_t)^2] \quad \text{since } u_t \text{ is uncorrelated with both } q_t \text{ and } Q_t \\ &= \sigma_u^2 + E[\lambda_0^2 q_t^2 + 2\lambda_0\lambda_1 q_t Q_t + \lambda_1^2 Q_t^2] \\ &= \sigma_u^2 + \lambda_0^2 + \lambda_1^2\sigma_Q^2 + 2\lambda_0\lambda_1\sqrt{\frac{2}{\pi}}\sigma_Q . \end{aligned}$$

To compute σ_w^2 using Equation 82 or $\sigma_w^2 = [\theta(1)]_1 \Omega [\theta(1)]_1'$ we need

$$\theta(L) = \left[\begin{array}{ccc} 1 & \lambda_0 + c_0 & \lambda_1 + c_1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] + \left[\begin{array}{ccc} 0 & -c_0 & -c_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] L \quad \text{so} \quad \theta(1) = \left[\begin{array}{ccc} 1 & \lambda_0 & \lambda_1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right].$$

Thus we get

$$\sigma_w^2 = \left[\begin{array}{ccc} 1 & \lambda_0 & \lambda_1 \end{array}\right]\left[\begin{array}{ccc} \sigma_u^2 & 0 & 0 \\ 0 & 1 & \sqrt{\frac{2}{\pi}}\sigma_Q \\ 0 & \sqrt{\frac{2}{\pi}}\sigma_Q & \sigma_Q^2 \end{array}\right]\left[\begin{array}{c} 1 \\ \lambda_0 \\ \lambda_1 \end{array}\right] = \sigma_u^2 + \lambda_0^2 + 2\sqrt{\frac{2}{\pi}}\sigma_Q\lambda_0\lambda_1 + \lambda_1^2\sigma_Q^2 ,$$

the same expression as before.

Part (c): How much of the total efficient price variation is due to q_t? This depends on the ordering of the variables. If they are ordered as Q_t, q_t, u_t then

$$\Omega_1 = \text{Var}\left(\left[\begin{array}{c} Q_t \\ q_t \\ u_t \end{array}\right]\right) = \left[\begin{array}{ccc} \sigma_Q^2 & \sqrt{\frac{2}{\pi}}\sigma_Q & 0 \\ \sqrt{\frac{2}{\pi}}\sigma_Q & 1 & 0 \\ 0 & 0 & \sigma_u^2 \end{array}\right].$$

We want to perform Cholesky decomposition of this matrix. The Cholesky decomposition of the 2×2 sub-block $\begin{bmatrix} \sigma_Q^2 & \sqrt{\frac{2}{\pi}}\sigma_Q \\ \sqrt{\frac{2}{\omega\pi}}\sigma_Q & 1 \end{bmatrix}$ is

$$F'F = \begin{bmatrix} \sigma_Q & 0 \\ \sqrt{\frac{2}{\pi}} & \sqrt{1-\frac{2}{\pi}} \end{bmatrix} \begin{bmatrix} \sigma_Q & \sqrt{\frac{2}{\pi}} \\ 0 & \sqrt{1-\frac{2}{\pi}} \end{bmatrix}.$$

Thus we can write σ_w^2 using Equation 82 in this new ordering (the same output ordering of $\begin{bmatrix} \Delta p_t \\ q_t \\ Q_t \end{bmatrix}$ we get that $\theta(L)$ given by

$$\theta(L) = \begin{bmatrix} \lambda_1 + c_1 & \lambda_0 + c_0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} -c_1 & -c_0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} L \quad \text{so} \quad [\theta(1)]_1 = \begin{bmatrix} \lambda_1 & \lambda_0 & 1 \end{bmatrix}.$$

Equation 82 would then give

$$\begin{aligned} [\theta(1)]_1 \Omega [\theta(1)]_1' &= \begin{bmatrix} \lambda_1 & \lambda_0 & 1 \end{bmatrix} \begin{bmatrix} \sigma_Q & 0 & 0 \\ \sqrt{\frac{2}{\pi}} & \sqrt{1-\frac{2}{\pi}} & 0 \\ 0 & 0 & \sigma_u \end{bmatrix} \begin{bmatrix} \sigma_Q & \sqrt{\frac{2}{\pi}} & 0 \\ 0 & \sqrt{1-\frac{2}{\pi}} & 0 \\ 0 & 0 & \sigma_u \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_0 \\ 1 \end{bmatrix} \\ &= \left(\lambda_1 \sigma_Q + \lambda_0 \sqrt{\frac{2}{\pi}} \right)^2 + \lambda_0 \left(1 - \frac{2}{\pi} \right) + \sigma_u^2. \end{aligned}$$

Thus the q_t variable 9the second one) contributes $\lambda_0 \left(1 - \frac{2}{\lambda_0}\right)$ to the total variance.

When the order of the the variables is $\begin{bmatrix} q_t \\ Q_t \\ u_t \end{bmatrix}$ we follow the same steps as earlier.

- Write $\Omega = \text{Var}\left(\begin{bmatrix} q_t \\ Q_t \\ u_t \end{bmatrix} \right)$.
- Compute its Cholesky factorization.
- Determine what $\theta(L)$ and $\theta(1)$ are in this new coordinate ordering.
- Express σ_w^2 in terms of the components of q_t, Q_t, and u_t.

Step one and two are

$$\Omega_{\left[\begin{array}{ccc} q_t & Q_t & u_t \end{array}\right]^T} = \mathrm{Var}\left(\left[\begin{array}{c} q_t \\ Q_t \\ u_t \end{array}\right]\right) = \left[\begin{array}{ccc} 1 & \sqrt{\frac{2}{\pi}}\sigma_Q & 0 \\ \sqrt{\frac{2}{\pi}} & \sigma_Q^2 & 0 \\ 0 & 0 & \sigma_u^2 \end{array}\right]$$

$$= \left[\begin{array}{ccc} 1 & 0 & 0 \\ \sqrt{\frac{2}{\pi}}\sigma_Q & \sqrt{\sigma_Q^2 - \frac{2}{\pi}\sigma_Q^2} & 0 \\ 0 & 0 & \sigma_u \end{array}\right] \left[\begin{array}{ccc} 1 & \sqrt{\frac{2}{\pi}}\sigma_Q & 0 \\ 0 & \sqrt{1 - \frac{2}{\pi}}\sigma_Q & 0 \\ 0 & 0 & \sigma_u \end{array}\right].$$

For step three, the expression for $\theta(L)$ with this variable ordering is

$$\theta(L) = \left[\begin{array}{ccc} \lambda_0 + c_0 & \lambda_1 + c_1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right] + \left[\begin{array}{ccc} -c_0 & -c_1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] L \quad \text{so} \quad [\theta(1)]_1 = \left[\begin{array}{ccc} \lambda_0 & \lambda_1 & 1 \end{array}\right].$$

Again under this ordering Equation 82 finally give for σ_w^2 the following

$$\sigma_w^2 = \left[\begin{array}{ccc} \lambda_0 & \lambda_1 & 1 \end{array}\right] \left[\begin{array}{ccc} 1 & 0 & 0 \\ \sqrt{\frac{2}{\pi}}\sigma_Q & \sqrt{1 - \frac{2}{\pi}}\sigma_Q & 0 \\ 0 & 0 & \sigma_u \end{array}\right] \left[\begin{array}{ccc} 1 & \sqrt{\frac{2}{\pi}}\sigma_Q & 0 \\ 0 & \sqrt{1 - \frac{2}{\pi}}\sigma_Q & 0 \\ 0 & 0 & \sigma_u \end{array}\right] \left[\begin{array}{c} \lambda_0 \\ \lambda_1 \\ 1 \end{array}\right]$$

$$= \left(\lambda_0 + \sqrt{\frac{2}{\pi}}\sigma_Q\lambda_1\right)^2 + \left(1 - \frac{2}{\pi}\right)\sigma_Q^2\lambda_1^2 + \sigma_u^2.$$

Thus the contribution to the total variance of σ_w^2 provided by the variable q_t (which is the first variable in this ordering) is given by $\left(\lambda_0 + \sqrt{\frac{2}{\pi}}\sigma_Q\lambda_1\right)^2$ as claimed.

Problem 9.2 (the model of Madhavan, Richardson, Roomans)

For the $y_t = \left[\begin{array}{c} \Delta p_t \\ q_t \end{array}\right]$ we have

$$\begin{aligned} \Delta p_t &= p_t - p_{t-1} = m_t + cq_t - m_{t-1} - cq_{t-1} = m_{t-1} + w_t + cq_t - m_{t-1} - cq_{t-1} \\ &= w_t + c(q_t - q_{t-1}) = u_t + \lambda v_t + c(\rho q_{t-1} + v_t - q_{t-1}) \\ &= u_t + \lambda v_t + c(\rho - 1)q_{t-1} + cv_t = u_t + (\lambda + c)v_t + c(\rho - 1)q_{t-1} \\ &= \left[\begin{array}{cc} 1 & \lambda + c \end{array}\right] \left[\begin{array}{c} u_t \\ v_t \end{array}\right] + \left[\begin{array}{cc} 0 & c(\rho - 1) \end{array}\right] \left[\begin{array}{c} \Delta p_{t-1} \\ q_{t-1} \end{array}\right]. \end{aligned}$$

For q_t we have $q_t = \rho q_{t-1} + v_t$ i.e. it is already in AR form. Thus our vector model is

$$\left[\begin{array}{c} \Delta p_t \\ q_t \end{array}\right] = \left[\begin{array}{cc} 1 & \lambda + c \\ 0 & 1 \end{array}\right] \left[\begin{array}{c} u_t \\ v_t \end{array}\right] + \left[\begin{array}{cc} 0 & c(\rho - 1) \\ 0 & \rho \end{array}\right] \left[\begin{array}{c} \Delta p_{t-1} \\ q_{t-1} \end{array}\right] \equiv \theta_0 \left[\begin{array}{c} u_t \\ v_t \end{array}\right] + \phi_1 \left[\begin{array}{c} \Delta p_{t-1} \\ q_{t-1} \end{array}\right],$$

where we have defined θ_0 and ϕ_0. Notice there are not other coefficient thus $\theta_1 = \phi_2 = 0$. This is a first order VAR process in the form $\phi(L)y_t = \theta(L)\varepsilon_t$. We can write this expression in the form

$$\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & c(\rho - 1) \\ 0 & \rho \end{bmatrix} L \right) \begin{bmatrix} \Delta p_t \\ q_t \end{bmatrix} = \begin{bmatrix} 1 & \lambda + c \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_t \\ v_t \end{bmatrix}.$$

For the given VAR model $\phi(L)y_t = \theta(L)\varepsilon_t$ we have $\sigma_w^2 = A\Omega A'$ with A the first row of the matrix product $(\phi(L))^{-1}\theta(1)$. For this problem the expression $\phi(1) = \begin{bmatrix} 1 & -c(\rho-1) \\ 0 & 1-\rho \end{bmatrix}$ so

$$\phi(1)^{-1}\theta(1) = \frac{1}{1-\rho} \begin{bmatrix} 1-\rho & c(\rho-1) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \lambda + c \\ 0 & 1 \end{bmatrix} = \frac{1}{1-\rho} \begin{bmatrix} 1-\rho & \lambda(1-\rho) \\ 0 & 1 \end{bmatrix}.$$

The first row is $\begin{bmatrix} 1 & \lambda \end{bmatrix}$ so

$$\sigma_w^2 = \begin{bmatrix} 1 & \lambda \end{bmatrix} \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix} \begin{bmatrix} 1 \\ \lambda \end{bmatrix} = \sigma_u^2 + \lambda^2 \sigma_v^2 .$$

We can compute this directly using the expression $w_t = u_t + \lambda v_t$ since u_t and v_t are independent and we get

$$\sigma_w^2 = \sigma_u^2 + \lambda^2 \sigma_v^2 ,$$

the same result.

Chapter 10 (Multiple Securities and Multiple Prices)

Notes on the text

Notes on the stacked models of multiple prices

For the given structural model we can compute Γ_0 as

$$\Gamma_0 = \text{Var}(\Delta p_t) = E\left[\left[\begin{array}{cc} \Delta p_{1,t} & \Delta p_{2,t} \end{array}\right]\left[\begin{array}{c} \Delta p_{1,t} \\ \Delta p_{2,t} \end{array}\right]\right] = E\left[\left[\begin{array}{cc} \Delta p_{1,t}{}^2 & \Delta p_{1,t}\Delta p_{2,t} \\ \Delta p_{1,t}\Delta p_{2,t} & \Delta p_{2,t}{}^2 \end{array}\right]\right]$$

In Equation 8 we have already computed the values of the diagonal of the above matrix. The off-diagonal elements can be computed using Equation 4 where we find

$$\begin{aligned} E[\Delta p_{1,t}\Delta p_{2,t}] &= E[(u_{1,t} + c(q_{1,t} - q_{1,t-1}))(u_{2,t} + c(q_{2,t} - q_{2,t-1}))] \\ &= c^2 E[(q_{1,t} - q_{1,t-1})(q_{2,t} - q_{2,t-1})] = c^2(E[q_{1,t}q_{2,t}] + E[q_{1,t-1}q_{2,t-1}]) \\ &= c^2(\rho + \rho) = 2\rho c^2 \,. \end{aligned}$$

This gives

$$\Gamma_0 = \left[\begin{array}{cc} \sigma_u^2 + 2c^2 & 2\rho c^2 \\ 2\rho c^2 & \sigma_u^2 + 2c^2 \end{array}\right] . \tag{87}$$

We can compute Γ_1 in the same way

$$\Gamma_1 = \text{Cov}(\Delta p_t, \Delta p_{t-1}) = E\left[\left[\begin{array}{cc} \Delta p_{1,t} & \Delta p_{2,t} \end{array}\right]\left[\begin{array}{c} \Delta p_{1,t-1} \\ \Delta p_{2,t-1} \end{array}\right]\right] = E\left[\left[\begin{array}{cc} \Delta p_{1,t}\Delta p_{1,t-1} & \Delta p_{1,t}\Delta p_{2,t-1} \\ \Delta p_{2,t}\Delta p_{1,t-1} & \Delta p_{2,t}\Delta p_{2,t-1}] \end{array}\right]\right] .$$

We can evaluate each element in in this matrix in tern

$$\begin{aligned} E[\Delta p_{1,t}\Delta p_{1,t-1}] &= E[(u_{1,t} + c(q_{1,t} - q_{1,t-1}))(u_{1,t-1} + c(q_{1,t-1} - q_{1,t-2}))] = -c^2 E[q_{1,t-1}^2] = -c^2 \\ E[\Delta p_{1,t}\Delta p_{2,t-1}] &= E[(u_{1,t} + c(q_{1,t} - q_{1,t-1}))(u_{2,t-1} + c(q_{2,t-1} - q_{2,t-2}))] = -c^2 E[q_{1,t-1}q_{2,t-1}] = -c^2\rho \\ E[\Delta p_{2,t}\Delta p_{1,t-1}] &= E[(u_{2,t} + c(q_{2,t} - q_{2,t-1}))(u_{1,t-1} + c(q_{1,t-1} - q_{1,t-2}))] = -c^2 E[q_{2,t-1}q_{1,t-1}] = -c^2\rho \,. \end{aligned}$$

This gives for the matrix Γ_1

$$\Gamma_1 = \left[\begin{array}{cc} -c^2 & -\rho c^2 \\ -\rho c^2 & c^2 \end{array}\right] , \tag{88}$$

the same as that in the book.

Notes on the suggested structural model

The model suggested in the book in this section is given by

$$\begin{aligned} m_t &= m_{t-1} + u_t \\ p_{1,t} &= m_t + cq_t \\ p_{2,t} &= m_{t-1} \,. \end{aligned}$$

For the vector Δp_t and this structural model we compute

$$\Delta p_t = \begin{bmatrix} \Delta p_{1,t} \\ \Delta p_{2,t} \end{bmatrix} = \begin{bmatrix} m_t + cq_t - m_{t-1} - cq_{t-1} \\ m_{t-1} - m_{t-2} \end{bmatrix} = \begin{bmatrix} u_t + cq_t - cq_{t-1} \\ u_{t-1} \end{bmatrix}$$
$$= \begin{bmatrix} 1 & c \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_t \\ q_t \end{bmatrix} + \begin{bmatrix} 0 & -c \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_{t-1} \\ q_{t-1} \end{bmatrix} .$$

This last expression defines the matrices θ_0^* and θ_1^* and the vector ε_t^*. This is one VMA representation of the vector Δp but one which cannot be transformed into a VAR model.

To compute a VMA model that can be converted into a VAR model, we first compute Γ_0 and Γ_1 from the vector Δp_t. From the definition of Γ_0 we find

$$\Gamma_0 = E\left[\begin{bmatrix} u_t + c(q_t - q_{t-1}) \\ u_{t-1} \end{bmatrix} \begin{bmatrix} u_t + c(q_t - q_{t-1}) & u_{t-1} \end{bmatrix}\right] = \begin{bmatrix} 2c^2 + \sigma_u^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix} , \tag{89}$$

using earlier results to evaluate the expectations. From the definition of Γ_1 we find

$$\Gamma_1 = E\left[\begin{bmatrix} u_t + c(q_t - q_{t-1}) \\ u_{t-1} \end{bmatrix} \begin{bmatrix} u_{t-1} + c(q_{t-1} - q_{t-2}) & u_{t-2} \end{bmatrix}\right] = \begin{bmatrix} -c^2 & 0 \\ \sigma_u^2 & 0 \end{bmatrix} . \tag{90}$$

As a second step, for a *general* VMA(1) process written in the form

$$\Delta p_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} ,$$

we can compute Γ_0 and Γ_1 (which will be functions of the unknown matrix θ) and then match these results to the explicit expressions for Γ_0 and Γ_1 computed above for the specific structural model we are considering here. These relationships will form a set of equations that we can solve to determine the elements of the matrix θ. For the VMA(1) model $\Delta p_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}$ for Γ_0 we find

$$\Gamma_0 = E[\Delta p_t \Delta p_t'] = E[(\varepsilon_t + \theta_1 \varepsilon_{t-1})(\varepsilon_t' + \varepsilon_{t-1}' \theta_1')] = E[\varepsilon_t \varepsilon_t] + \theta_1 E[\varepsilon_{t-1} \varepsilon_{t-1}'] \theta_1'$$
$$= \Omega + \theta_1 \Omega \theta_1' , \tag{91}$$

and for Γ_1 we find

$$\Gamma_1 = E[\Delta p_t \Delta p_{t-1}'] = E[(\varepsilon_t + \theta_1 \varepsilon_{t-1})(\varepsilon_{t-1}' + \varepsilon_{t-2}' \theta_1')]$$
$$= \theta_1 \Omega . \tag{92}$$

Since we know Γ_0 and Γ_1 from our structural model via Equations 89 and 90 we can attempt to find a matrix of θ in the general VMA model such that the general model has the same values of Γ_0 and Γ_1. Using Equation 91 and 92 this means that

$$\Omega + \theta_1 \Omega \theta_1' = \Gamma_0 \tag{93}$$
$$\theta_1 \Omega = \Gamma_1 . \tag{94}$$

These are two equations for the two "variables" θ_1 and Ω. In general we don't know that the matrix θ_1 is invertible and in fact for the structural model here it is not invertible. We would expect that Ω is invertible however. Thus from Equation 94 we would have $\theta_1 = \Gamma_1 \Omega^{-1}$, which when we put this into Equation 93 we would get

$$\Omega + \Gamma_1 (\Omega^{-1})' \Gamma_1' = \Gamma_0 .$$

or one equation for Ω. Using the Mathematica script `chap_10_the_VMA_representation.nb` we can evaluate each side of this expression and then solve for the elements of the matrix Ω. We find

$$\Omega = \frac{1}{c^2+\sigma_u^2} \left[\begin{array}{cc} c^4 + 3c^2\sigma_u^2 + \sigma_u^4 & c^2\sigma_u^2 \\ c^2\sigma_u^2 & c^2\sigma_u^2 \end{array} \right] \quad (95)$$

Once we have Ω we can compute $\theta_1 = \Gamma_1 \Omega^{-1}$ where we get

$$\theta_1 = \frac{1}{c^2+\sigma_u^2} \left[\begin{array}{cc} -c^2 & c^2 \\ \sigma_u^2 & -\sigma_u^2 \end{array} \right] . \quad (96)$$

Note that the above expression for Ω and θ_1 are equivalent to those presented in the book when we simplify some.

With the filtered estimate of f_t given by $f_t = p_t + \theta_1 \varepsilon_t$ we have that the first difference of f_t given by

$$\begin{aligned} \Delta f_t &= \Delta p_t + \theta_1 \Delta \varepsilon_t \quad \text{and using the general VMA(1) model for } \Delta p_t \\ &= \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_1(\varepsilon_t - \varepsilon_{t-1}) \\ &= (I + \theta_1)\varepsilon_t \, . \end{aligned} \quad (97)$$

Exercise 10.1 (the innovations in σ_u^2)

If the ordering of the prices is $\left[\begin{array}{cc} p_{1,t} & p_{2,t} \end{array} \right]$ then from earlier results in this chapter to explain the decomposition of σ_u^2 we would first need to compute θ_1 and Ω. In the python code `chap_10_structural_model.py` using the numbers given we first compute

$$\theta_1 = \left[\begin{array}{cc} -0.8 & 0.8 \\ 0.2 & -0.2 \end{array} \right] \quad \text{and} \quad \Omega = \left[\begin{array}{cc} 5.8 & 0.8 \\ 0.8 & 0.8 \end{array} \right] .$$

The Cholesky factor of $\Omega = F'F$ has $F' = \left[\begin{array}{cc} 2.408 & 0. \\ 0.3321 & 0.83045 \end{array} \right]$ and we then compute

$$\begin{aligned} \sigma_u^2 &= ([\theta(1)]_1 F')([\theta(1)]_1 F)' = \left[\begin{array}{cc} 0.74740932 & 0.66436384 \end{array} \right] \left[\begin{array}{c} 0.74740932 \\ 0.66436384 \end{array} \right] \\ &= 0.55862069 + 0.44137931 = 1.0 \, . \end{aligned}$$

Thus under this model the innovation from the first price explains 55.9% of σ_u^2 while the innovation from the second price explains 44.1% of σ_u^2.

In the *alternative* ordering for the prices our model for the vector price change $\Delta \hat{p}_t$ is

$$\Delta \hat{p}_t = \left[\begin{array}{c} \Delta p_{t,2} \\ \Delta p_{t,1} \end{array} \right] = \left[\begin{array}{c} u_{t-1} \\ u_t + c(q_t - q_{t-1}) \end{array} \right] .$$

From this we find the autocovariance of $\Delta \hat{p}_t$ are

$$\hat{\Gamma}_0 = \left[\begin{array}{cc} \sigma_u^2 & 0 \\ 0 & 2c^2 + \sigma_u^2 \end{array} \right] ,$$

and

$$\hat{\Gamma}_1 = E\left[\left[\begin{array}{c} u_{t-1} \\ u_t + c(q_t - q_{t-1}) \end{array}\right]\left[\begin{array}{cc} u_{t-2} & u_{t-1} + c(q_{t-1} - q_{t-2}) \end{array}\right]\right] = \left[\begin{array}{cc} 0 & \sigma_u^2 \\ 0 & -c^2 \end{array}\right].$$

Now we have to compute a VMA(1) representative for this ordering of prices of the form $\Delta\hat{p}_t = \hat{\varepsilon}_t + \hat{\theta}_1\hat{\varepsilon}_{t-1}$ that has the same $\hat{\Gamma}_0$ and $\hat{\Gamma}_1$ as we computed above. This means solving Equations 93 and 94 for $\hat{\theta}_1$ and $\hat{\Omega}$. Following the same steps from before we have that

$$\hat{\Omega} = \frac{1}{c^2 + \sigma_u^2}\left[\begin{array}{cc} c^2\sigma_u^2 & c^2\sigma_u^2 \\ c^2\sigma_u^2 & c^4 + 3c^2\sigma_u^2 + \sigma_u^4 \end{array}\right],$$

and

$$\hat{\theta}_1 = \frac{1}{c^2 + \sigma_u^2}\left[\begin{array}{cc} -\sigma_u^2 & \sigma_u^2 \\ c^2 & -c^2 \end{array}\right].$$

Following the same steps as before this then gives

$$\sigma_u^2 = 0.8 + 0.2 = 1.0\,.$$

Remembering that the second term in the above sum is the contribution to σ_u^2 from the first price we see that it is 20% as expected.

Notes on the autoregressive representation

Inverting $I + \theta_1 L$ to get $\varphi(L)$ we have

$$\begin{aligned}
(I + \theta_1 L)^{-1} &= I - \theta_1 L + \theta_1^2 L^2 - \theta_1^3 L^3 + \theta_1^4 L^4 - \cdots \\
&= I - \frac{1}{c^2 + \sigma_u^2}\left[\begin{array}{cc} -c^2 & c^2 \\ \sigma_u^2 & -\sigma_u^2 \end{array}\right] L + \frac{1}{c^2 + \sigma_u^2}\left[\begin{array}{cc} c^2 & -c^2 \\ -\sigma_u^2 & \sigma_u^2 \end{array}\right] L^2 \\
&- \frac{1}{c^2 + \sigma_u^2}\left[\begin{array}{cc} -c^2 & c^2 \\ \sigma_u^2 & -\sigma_u^2 \end{array}\right] L^3 + \frac{1}{c^2 + \sigma_u^2}\left[\begin{array}{cc} c^2 & -c^2 \\ -\sigma_u^2 & \sigma_u^2 \end{array}\right] L^4 + \cdots
\end{aligned}$$

when we compute the needed matrix powers of θ_1 given by Equation 96.

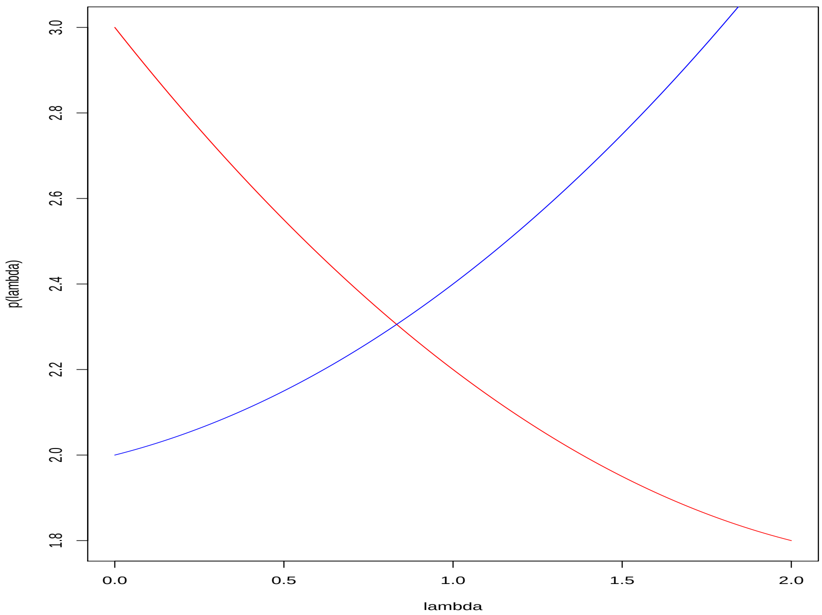

Figure 1: The arrival intensity for Exercise 11.1. The red curve is the function $\lambda^{\text{Buy}}(p)$ and the blue curve is the function $\lambda^{\text{Sell}}(p)$.

Chapter 11 (Dealers and Their Inventories)

Notes on the text

Exercise 11.1 (the optimal price based on arrival rates)

This problem is worked in the R code ex_11_1.R. The equilibrium price and rate are the price and rate where the two arrival rate curves $\lambda^{\text{Buy}}(p)$ and $\lambda^{\text{Sell}}(p)$ intersect. The dealers average profit (or trading revenue) per unit time is given by the function π or

$$\pi(\text{Bid}, \text{Ask}) = (\text{Ask} - \text{Bid})\lambda^{\text{Buy}}(\text{Ask}) = (\text{Ask} - \text{Bid})\lambda^{\text{Sell}}(\text{Bid})\,. \quad (98)$$

Since in this problem we are given the prices as a function of the arrival rate λ we should write the above as

$$\pi = (p^{\text{Buy}}(\lambda) - p^{\text{Sell}}(\lambda))\lambda\,.$$

We then seek to find the value of λ that maximizes this function. Doing this numerically we find

```
[1] "max profit= 0.208264; lambda= 0.424242; p_bid =   2.12; p_ask=   2.61"
```

Notes on risk aversion and dealer behavior)

In this section the book introduces a utility function given by $U(W) = -e^{-\alpha W}$ where W is a random variable with a normal distribution $W \sim \mathcal{N}(\mu_W, \sigma_W^2)$. Given this setup we can find the expectation of $U(W)$ using the definition of expectation as

$$
\begin{aligned}
E[U(W)] &= \int_{-\infty}^{\infty} (-e^{-\alpha W}) \left(\frac{1}{\sqrt{2\pi}\sigma_W} e^{-\frac{1}{2}\frac{(W-\mu_W)^2}{\sigma_W^2}} \right) dW \\
&= -\frac{1}{\sqrt{2\pi}\sigma_W} \int_{-\infty}^{\infty} \exp\left\{ -\alpha W - \frac{1}{2\sigma_W^2}(W^2 - 2\mu_W W + \mu_W^2) \right\} dW \\
&= -\frac{e^{-\frac{\mu_W^2}{2\sigma_W^2}}}{\sqrt{2\pi}\sigma_W} \int_{-\infty}^{\infty} \exp\left\{ -\alpha W - \frac{1}{2\sigma_W^2} W^2 + \frac{\mu_W}{\sigma_W^2} W \right\} dW \\
&= -\frac{e^{-\frac{\mu_W^2}{2\sigma_W^2}}}{\sqrt{2\pi}\sigma_W} \int_{-\infty}^{\infty} \exp\left\{ -\frac{1}{2\sigma_W^2} \left(W^2 - 2(\mu_W - \sigma_W^2 \alpha) W \right) \right\} dW \\
&= -\frac{e^{-\frac{\mu_W^2}{2\sigma_W^2}}}{\sqrt{2\pi}\sigma_W} \int_{-\infty}^{\infty} \exp\left\{ -\frac{1}{2\sigma_W^2} \left[(W - (\mu_W - \sigma_W^2\alpha))^2 - (\mu_W - \sigma_W^2 \alpha)^2 \right] \right\} dW \\
&= -\frac{e^{-\frac{\mu_W^2}{2\sigma_W^2} + \frac{(\mu_W - \sigma_W^2\alpha)^2}{2\sigma_W^2}}}{\sqrt{2\pi}\sigma_W} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma_W^2} W^2} dW \\
&= -\exp\left\{ -\frac{\mu_W^2}{2\sigma_W^2} + \frac{(\mu_W - \sigma_W^2 \alpha)^2}{2\sigma_W^2} \right\} = -\exp\left\{ -\alpha\mu_W + \frac{1}{2}\sigma_W^2 \alpha^2 \right\} , \qquad (99)
\end{aligned}
$$

when we simplify the argument. Thus to maximize the functional form for $E[U(W)]$ we can minimize the argument of the exponential or maximize its negative value which is

$$
\alpha \mu_W - \frac{1}{2}\alpha^2 \sigma_W^2 .
$$

Since α is a positive constant we could divide by it and the resulting expression is called **certainty equivalent** or **CE** given by

$$
CE(\mu_W, \sigma_W^2) \equiv \mu_W - \frac{1}{2}\alpha\sigma_W^2 . \qquad (100)
$$

Note that for a general portfolio the CE is a function of its two arguments the portfolio's mean μ_W and variance σ_W^2.

To the dealer to be indifferent as to whether his bid gets hit we must have CE equal in each of these cases. This gives

$$
CE(n\mu_X, n^2\sigma_X^2) = CE((n+1)\mu_X - B, (n+1)^2\sigma_X^2) .
$$

Using Equation 100 for each side this gives

$$
n\mu_X - \frac{\alpha}{2} n^2 \sigma_X^2 = (n+1)\mu_X - B - \frac{\alpha}{2}(n+1)^2 \sigma_X^2 .
$$

When we solve for B in the above we get

$$B = \mu_X - (2n+1)\frac{\alpha}{2}\sigma_X^2 \,. \tag{101}$$

If our portfolio initially starts out as $W = n(X - P)$ then via Equation 100 the certainty equivalent of this portfolio (P is constant) is given by

$$CE(\mu_W, \sigma_W^2) = CE(n(\mu_X - P), n^2\sigma_X^2) = n(\mu_X - P) - \frac{\alpha}{2}n^2\sigma_X^2 \,.$$

To maximize the CE as a function of n we take the first derivative and set the result equal to zero, and solve for n. We need to solve

$$\mu_X - P - \alpha n \sigma_X^2 = 0 \quad \text{or} \quad n = \frac{\mu_X - P}{\alpha\sigma_X^2} \,.$$

When we put this value of n into Equation 101 we get

$$\begin{aligned} B &= \mu_X - \left(\frac{2}{\alpha\sigma_X^2}(\mu_X - P) + 1\right)\frac{\alpha}{2}\sigma_X^2 \\ &= \mu_X - (\mu_X - P) - \frac{\alpha}{2}\sigma_X^2 = P - \frac{\alpha}{2}\sigma_X^2 \,, \end{aligned}$$

the same value the book gives.

To demonstrate extensions of the above ideas to a multidimensional case we consider for simplification two securities (denoted "one" and "two") of which our number of shares in each are held in the vector $n = \begin{bmatrix} n_1 & n_2 \end{bmatrix}$. With no trades, our terminal wealth in this portfolio is given by $W = n'X$ here X is a random vector of prices. This portfolio has a mean value of $\mu_W = n'\mu_X$ and a variance $\sigma_W^2 = n'\Omega n$. This gives a certainty equivalent of

$$\begin{bmatrix} n_1 & n_2 \end{bmatrix} X - \frac{\alpha}{2}\left(\begin{bmatrix} n_1 & n_2 \end{bmatrix}\Omega\begin{bmatrix} n_1 & n_2 \end{bmatrix}'\right) .$$

If our bid gets hit in the first security we buy one unit at B_1 and our wealth in this case is

$$W = \begin{bmatrix} n_1 + 1 & n_2 \end{bmatrix} X - B_1 \,.$$

This last portfolio has a mean and variance given by

$$\mu_W = \begin{bmatrix} n_1 + 1 & n_2 \end{bmatrix}\mu_X - B_1 \quad \text{and} \quad \sigma_W^2 = \begin{bmatrix} n_1 + 1 & n_2 \end{bmatrix}\Omega\begin{bmatrix} n_1 + 1 & n_2 \end{bmatrix}' .$$

This gives a certainty equivalent in this case of

$$\begin{aligned} CE &= \begin{bmatrix} n_1 + 1 & n_2 \end{bmatrix} X - B_1 - \frac{\alpha}{2}\left(\begin{bmatrix} n_1 + 1 & n_2 \end{bmatrix}\Omega\begin{bmatrix} n_1 + 1 & n_2 \end{bmatrix}'\right) \\ &= \begin{bmatrix} n_1 & n_2 \end{bmatrix} X + \begin{bmatrix} 1 & 0 \end{bmatrix} X - B_1 \\ &- \frac{\alpha}{2}\left(\begin{bmatrix} n_1 & n_2 \end{bmatrix}\Omega\begin{bmatrix} n_1 & n_2 \end{bmatrix}' + \begin{bmatrix} 1 & 0 \end{bmatrix}\Omega\begin{bmatrix} n_1 & n_2 \end{bmatrix}' \right. \\ &+ \left. \begin{bmatrix} n_1 & n_2 \end{bmatrix}\Omega\begin{bmatrix} 1 & 0 \end{bmatrix}' + \begin{bmatrix} 1 & 0 \end{bmatrix}\Omega\begin{bmatrix} 1 & 0 \end{bmatrix}'\right) . \end{aligned}$$

When we equate these two expressions for the certainty equivalent and solve for B_1 we find

$$B_1 = \begin{bmatrix} 1 & 0 \end{bmatrix} \mu_X - \frac{\alpha}{2}\left(2\begin{bmatrix} 1 & 0 \end{bmatrix}\Omega\begin{bmatrix} n_1 & n_2 \end{bmatrix}' + \begin{bmatrix} 1 & 0 \end{bmatrix}\Omega\begin{bmatrix} 1 & 0 \end{bmatrix}'\right).$$

If Ω is given by $\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$ then the inner products in the above expression are computed as

$$\begin{aligned}
\begin{bmatrix} 1 & 0 \end{bmatrix}\Omega\begin{bmatrix} n_1 & n_2 \end{bmatrix}' &= \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \end{bmatrix}\begin{bmatrix} n_1 \\ n_2 \end{bmatrix} = \sigma_1^2 n_1 + \rho\sigma_1\sigma_2 n_2 \\
\begin{bmatrix} 1 & 0 \end{bmatrix}\Omega\begin{bmatrix} 1 & 0 \end{bmatrix}' &= \sigma_1^2 .
\end{aligned}$$

Thus we get for B_1

$$\begin{aligned}
B_1 &= \mu_1 - \frac{\alpha}{2}\left(2(\sigma_1^2 n_1 + \rho\sigma_1\sigma_2 n_2) + \sigma_1^2\right) \\
&= \mu_1 - \frac{\alpha\sigma_1}{2}\left((2n_1 + 1)\sigma_1 + 2\rho n_2\sigma_2\right) ,
\end{aligned} \tag{102}$$

the same expression in the book.

If we next consider the case where the dealer can trade some amount of the first security at a prices P (here a vector) then the portfolio is given by $W = \begin{bmatrix} n_1 & n_2 \end{bmatrix}(X - P)$ and we have a certainty equivalent given by

$$\begin{bmatrix} n_1 & n_2 \end{bmatrix}(\mu_X - P) - \frac{\alpha}{2}\left(\begin{bmatrix} n_1 & n_2 \end{bmatrix}\Omega\begin{bmatrix} n_1 & n_2 \end{bmatrix}'\right).$$

To minimize this with respect to the vector $\begin{bmatrix} n_1 & n_2 \end{bmatrix}$ we take the derivative to get

$$\mu_X - P - \alpha\Omega\begin{bmatrix} n_1 \\ n_2 \end{bmatrix}.$$

Setting this equal to zero and solving we get

$$\begin{aligned}
\begin{bmatrix} n_1 \\ n_2 \end{bmatrix} &= \frac{1}{\alpha}\Omega^{-1}(\mu_X - P) \\
&= \frac{1}{\alpha}\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}^{-1}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} - \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}\right) \\
&= \frac{1}{\alpha(\sigma_1^2\sigma_2^2(1 - \rho^2))}\begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix}\begin{bmatrix} \mu_1 - P_1 \\ \mu_2 - P_2 \end{bmatrix}.
\end{aligned}$$

If we put the expressions for n_1 and n_2 just computed into Equation 102 (in the Mathematical file `simplify_dealer_starting_at_optimum.nb`) we find

$$B_1 = 2P_1 - \mu_1 - \frac{\alpha}{2}\sigma_1^2 .$$

Note that this is different than the expression that the book has. If anyone sees anything wrong with what I have done please contact me.

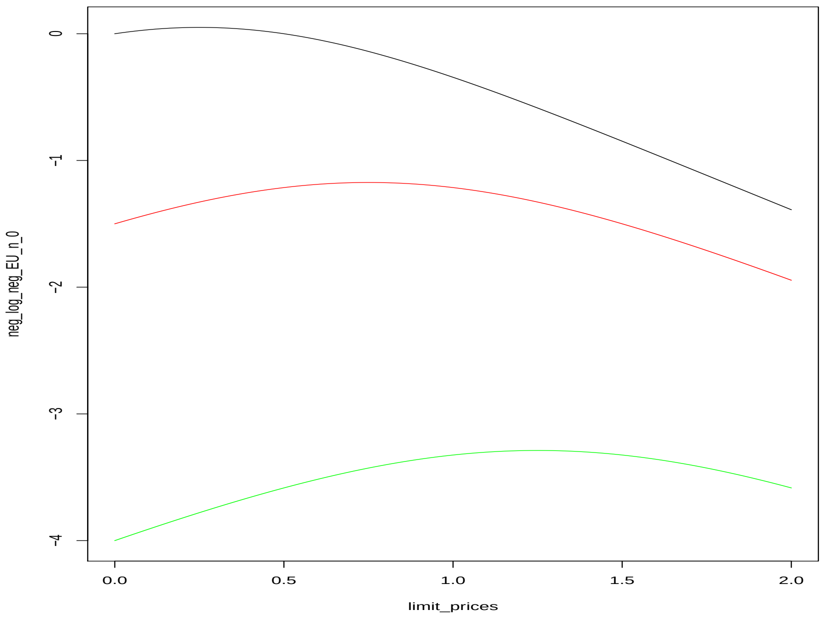

Figure 2: A duplication of Figure 12.1 from Chapter 12 in the book for $n = 0, -1, -2$. The maximization of the expected utility depends on the number of shares currently held n.

Chapter 12 (Limit Order Markets)

Notes on the text

Notes on the choice between a limit and market order

With the definitions given and assumed in this chapter for a limit order placed at L an exponential utility of $U(W) = -e^{-\alpha W}$, and the normal distribution approximation for wealth distribution W we can derive the expected utility given that the limit order might or might (hit) not get executed (not hit or the base case) to get

$$\begin{aligned}\mathrm{EU}_{\mathrm{Limit}}(L) &= \mathrm{Pr}_{\mathrm{Hit}}(L)\mathrm{EU}_{\mathrm{Hit}}(L) + (1 - \mathrm{Pr}_{\mathrm{Hit}}(L))\mathrm{EU}_{\mathrm{Base}} \\ &= (1 - e^{-\lambda(L-\theta)})\left(-e^{-\alpha((n+1)\mu_X - L) + \frac{\alpha^2}{2}(n+1)^2\sigma_X^2}\right) + e^{-\lambda(L-\theta)}\left(-e^{-\alpha n\mu_X + \frac{\alpha^2}{2}n^2\sigma_X^2}\right).\end{aligned}$$

For the the specific numbers given in the book $\alpha = 1$, $\mu_X = 1$, etc. we get for $\mathrm{EU}_{\mathrm{Limit}}(L)$ the following expression

$$\mathrm{EU}_{\mathrm{Limit}}(L) = (1 - e^{-L})(-e^{-(n+1-L)+\frac{1}{2}(n+1)^2}) + e^{-L}(-e^{-n+\frac{1}{2}n^2}).$$

We can take the negative of this expression, the logarithm, and the negative again and plot this for $n = 0$, $n = -1$, and $n = -2$. When we do that in `chap_12_plot_EU.R` we get the plot given in Figure 2. This plot matches well with the one given in the book.

If the customer enters a market order to buy then his expected utility is given by

$$\text{EU}_{\text{Market}} = -e^{-(n+1-A)+\frac{1}{2}(n+1)^2} .$$

We set this equal to $\text{EU}_{\text{Base}} = -e^{-n+\frac{1}{2}n^2}$ or the expected utility of doing nothing and solve for A to determine when the market order is preferred. We find that there is no preference between a market order and doing nothing when $A = -n + \frac{1}{2}$. Thus if $n = -1$ we get $A = \frac{3}{2}$. If $A < \frac{3}{2}$ the customer would prefer a market order and if $A > \frac{3}{2}$ the customer would prefer to do nothing.

Chapter 13 (Depth)

Notes on the text

Notes on the customer orders in the CARA-normal framework

Recall that $R(q)$ is the required capital expenditure to purchase q shares of our stock. The reason it is a function of q is that the customer might have to trade into the book at poorer prices to get all the desired shares. If the customer takes these trades his terminal wealth is given by

$$W = (n+q)X - R(q)\,,$$

since he ends the transaction with q more share of stock (with terminal value X) and pays $R(q)$ to get them. For the customer to act optimally he must seek to maximize his expected utility of his final wealth W or

$$\text{EU}(W) = E[-e^{-\rho W}]\,.$$

For an expected utility of this form we must *maximize* the certainty equivalent $CE(\mu_W, \sigma_W^2)$, where using Equation 100 is given by

$$CE(\mu_W, \sigma_W^2) \equiv \mu_W - \frac{1}{2}\rho\sigma_W^2 = ((n+q)\mu_X - R(q)) - \frac{1}{2}\rho(n+q)^2\sigma_X^2\,.$$

The optimal customer will want to maximize this with respect to q. To do this, we take the derivative with respect to q and set the result equal to zero to get

$$\mu_X - R'(q) - \rho(n+q)\sigma_X^2 = 0\,. \tag{103}$$

If the customer has a noisy signal that indicates his belief in the final price say $S = X + \varepsilon$ with $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. Then using Equation 34 since $\sigma_S^2 = \sigma_X^2 + \sigma_\varepsilon^2$ and $\sigma_{XS} = \sigma_X^2$ we have that

$$\mu_{X|S} = \mu_X + \frac{\sigma_{XS}}{\sigma_S^2}(S - \mu_X) = \mu_X + \frac{\sigma_X^2}{\sigma_X^2 + \sigma_\varepsilon^2}(S - \mu_X) = \frac{\mu_X\sigma_\varepsilon^2 + S\sigma_X^2}{\sigma_X^2 + \sigma_\varepsilon^2}\,. \tag{104}$$

Using Equation 36 we have that $\sigma_{X|S}^2$ is given by

$$\sigma_{X|S}^2 = \sigma_X^2 - \frac{\sigma_{XS}^2}{\sigma_S^2} = \sigma_X^2 - \frac{\sigma_X^4}{\sigma_X^2 + \sigma_\varepsilon^2} = \frac{\sigma_X^2\sigma_\varepsilon^2}{\sigma_X^2 + \sigma_\varepsilon^2}\,. \tag{105}$$

By moving terms across the equal sign we can write Equation 103 with $M \equiv R'(q)$ as

$$M + q\rho\sigma_X^2 = \mu_X - n\rho\sigma_X^2\,.$$

In the case that the customer thinks they have information on the final stock price X (from their signal S) we replace the above mean and variances with the conditional expressions to get

$$M + q\rho\sigma_{X|S}^2 = \mu_{X|S} - n\rho\sigma_{X|S}^2\,. \tag{106}$$

Denote the right-hand-side of this as ω. Then replacing $\mu_{X|S}$ and $\sigma^2_{X|S}$ with what they are given by from Equations 104 and 105 we get

$$\begin{aligned}\omega &= \frac{\mu_X \sigma_\varepsilon^2 + S\sigma_X^2}{\sigma_\varepsilon^2 + \sigma_X^2} - n\rho \frac{\sigma_\varepsilon^2 \sigma_X^2}{\sigma_\varepsilon^2 + \sigma_X^2} \\ &= \frac{\mu_X \sigma_\varepsilon^2 + \sigma_X^2 (S - \rho n \sigma_\varepsilon^2)}{\sigma_\varepsilon^2 + \sigma_X^2} = \frac{\mu_X \sigma_\varepsilon^2 + \sigma_X^2 (X + \varepsilon - \rho n \sigma_\varepsilon^2)}{\sigma_\varepsilon^2 + \sigma_X^2} \,. \end{aligned} \tag{107}$$

In the above expression X, ε, and n are random variables. Thus the measurement of ω should give us information on the value of the variables and in particular on X. To use Equations 34 and 36 we take expectation of both sides to get

$$\mu_\omega = E[\omega] = \frac{\sigma_\varepsilon^2 \mu_X + \sigma_X^2 (\mu_X - \sigma_\varepsilon^2 E[n]\rho)}{\sigma_\varepsilon^2 + \sigma_X^2} = \mu_X \,,$$

since $E[n] = 0$. Next we compute σ_ω^2 to find

$$\sigma_\omega^2 = \frac{\sigma_X^4}{(\sigma_\varepsilon^2 + \sigma_X^2)^2} \mathrm{Var}\left(X + \varepsilon - \sigma_\varepsilon^2 \rho n\right) = \frac{\sigma_X^4}{(\sigma_\varepsilon^2 + \sigma_X^2)^2} \left(\sigma_X^2 + \sigma_\varepsilon^2 + \sigma_\varepsilon^4 \rho^2 \sigma_n^2\right) \,. \tag{108}$$

Finally, we compute $\sigma_{\omega,X}$ where we find

$$\begin{aligned}\sigma_{\omega,X} &= E[(\omega - \mu_\omega)(X - \mu_X)] \\ &= E\left[\left(\frac{\mu_X \sigma_\varepsilon^2 + \sigma_X^2 (X + \varepsilon - \rho n \sigma_\varepsilon^2) - \mu_X(\sigma_\varepsilon^2 + \sigma_X^2)}{\sigma_\varepsilon^2 + \sigma_X^2}\right)(X - \mu_X)\right] \\ &= \frac{1}{\sigma_\varepsilon^2 + \sigma_X^2} E\left[(\sigma_X^2 (X - \mu_X) + \sigma_X^2 \varepsilon - \rho n \sigma_X^2 \sigma_\varepsilon^2)(X - \mu_X)\right] = \frac{\sigma_X^4}{\sigma_\varepsilon^2 + \sigma_X^2} \,. \end{aligned} \tag{109}$$

Using these two expressions we can now compute the expected value of X given ω from Equations 34 as

$$E[X|\omega] = \mu_X + \frac{\sigma_{\omega,X}}{\sigma_\omega^2}(\sigma - \mu_X) \,.$$

Notes on the competitive dealer market

The equilibrium condition $P(q) = \mu_{X|\omega}$ becomes

$$\begin{aligned} k_0 + k_1 q &= \mu_X + \frac{\sigma_{\omega,X}}{\sigma_\omega^2}(\omega - \mu_X) \\ &= \mu_X + \frac{\sigma_{\omega,X}}{\sigma_\omega^2}\left((k_0 + 2k_1 q) + q\rho\sigma_{X|S}^2 - \mu_X\right) \\ &= \mu_X - \frac{\sigma_{\omega,X}}{\sigma_\omega^2}(k_0 - \mu_X) + \frac{\sigma_{\omega,X}}{\sigma_\omega^2}(2k_1 + \rho\sigma_{X|S}^2)q \,. \end{aligned} \tag{110}$$

Equating powers of q on both sides we have

$$k_0 = \mu_X - \frac{\sigma_{\omega,X}}{\sigma_\omega^2}(k_0 - \mu_X) \quad \Rightarrow \quad k_0 = \mu_X \,, \tag{111}$$

and

$$k_1 = \frac{\sigma_{\omega,X}}{\sigma_\omega^2}(2k_1 + \rho + \sigma_{X|S}^2) \quad \Rightarrow \quad k_1 = \frac{\rho\sigma_{X|S}^2\sigma_{\omega,X}}{\sigma_\omega^2 - 2\sigma_{\omega,X}} = \frac{\rho\sigma_X^2\sigma_\varepsilon^2}{\rho^2\sigma_n^2\sigma_\varepsilon^4 - \sigma_X^2 - \sigma_\varepsilon^2}\,, \tag{112}$$

when we use Equations 105, 109, and 108.

For the price schedule $P(q)$ to be upward sloping i.e. $P'(q) > 0$ we must have $k_1 > 0$ or the denominator of Equation 112 nonnegative. Thus

$$\rho^2\sigma_n^2\sigma_\varepsilon^2 - \sigma_X^2 - \sigma_\varepsilon^2 > 0\,.$$

Using this inequality (if needed) and Equation 112 we can derive other relationships on k_1 by taking the derivatives suggested in the book. We find

- k_1 increases with σ_X^2 since

  $$\frac{\partial k_1}{\partial \sigma_X^2} = \frac{\rho\sigma_\varepsilon^2}{\rho^2\sigma_n^2\sigma_\varepsilon^4 - \sigma_X^2 - \sigma_\varepsilon^2} + \frac{\rho\sigma_X^2\sigma_\varepsilon^2}{(\rho^2\sigma_n^2\sigma_\varepsilon^4 - \sigma_X^2 - \sigma_\varepsilon^2)^2} > 0\,,$$

  since both expressions added above are positive.

- k_1 decreases with σ_ε^2 since

  $$\begin{aligned}\frac{\partial k_1}{\partial \sigma_\varepsilon^2} &= \frac{\rho\sigma_X^2}{\rho^2\sigma_n^2\sigma_\varepsilon^4 - \sigma_X^2 - \sigma_\varepsilon^2} - \frac{\rho\sigma_X^2\sigma_\varepsilon^2(2\rho^2\sigma_n^2\sigma_\varepsilon^2)}{(\rho^2\sigma_n^2\sigma_\varepsilon^4 - \sigma_X^2 - \sigma_\varepsilon^2)^2} + \frac{\rho\sigma_X^2\sigma_\varepsilon^2}{(\rho^2\sigma_n^2\sigma_\varepsilon^4 - \sigma_X^2 - \sigma_\varepsilon^2)^2} \\ &= -\frac{\rho^3\sigma_X^2\sigma_n^2\sigma_\varepsilon^4}{(\rho^2\sigma_n^2\sigma_\varepsilon^4 - \sigma_X^2 - \sigma_\varepsilon^2)^2}\,,\end{aligned}$$

  when we simplify.

- k_1 decreases with σ_n^2 since

  $$\frac{\partial k_1}{\partial \sigma_n^2} = -\frac{\rho^3\sigma_X^2\sigma_\varepsilon^6}{(\rho^2\sigma_n^2\sigma_\varepsilon^4 - \sigma_X^2 - \sigma_\varepsilon^2)^2} < 0\,.$$

- k_1 decreases with ρ since

  $$\begin{aligned}\frac{\partial k_1}{\partial \rho} &= \frac{\sigma_X^2\sigma_\varepsilon^2}{\rho^2\sigma_n^2\sigma_\varepsilon^4 - \sigma_X^2 - \sigma_\varepsilon^2} - \frac{\rho\sigma_X^2\sigma_\varepsilon^2(2\rho\sigma_n^2\sigma_\varepsilon^2)}{(\rho^2\sigma_n^2\sigma_\varepsilon^4 - \sigma_X^2 - \sigma_\varepsilon^2)^2} \\ &= \frac{-\rho^2\sigma_X^2\sigma_n^2\sigma_\varepsilon^4 - \sigma_X^4\sigma_\varepsilon^2 - \sigma_X^2\sigma_\varepsilon^4}{(\rho^2\sigma_n^2\sigma_\varepsilon^4 - \sigma_X^2 - \sigma_\varepsilon^2)^2} < 0\,,\end{aligned}$$

  since every term in the above sum is negative.

Chapter 14 (Trading Costs: Retrospective and Comparative)

Notes on the text

Notes on the implementation shortfall

In the definition of the implementation shortfall we assume that we currently hold an existing portfolio (this could be all cash) and desire to obtain a different portfolio for some reason. The desired portfolio is then expressed to a broker who will attempt to obtain that portfolio in a certain manner. The person or company tasked with this job may or may not succeed in getting all of the shares desired because the order was so large that it exhausted all of the shares available. In fact, they may get the desired shares but due to significant market impact have poor execution prices. The implementation shortfall is designed to measure the loss associated with not getting the complete desired portfolio and how poor the price paid was in terms of some standard prices. As such, we define several variables

- n_0 a vector of the initial portfolio holdings (in shares). The first component of which is the cash dollar amount.
- π_0 is a vector of initial "benchmark" prices of each of the stocks in our universe.
- v is a vector of the desired position (in shares). Again the first component of this vector represents cash.
- π_1 is a vector of final "benchmark" prices. Notionally, these are prices that we hope we can execute our trades at. Executions at prices worse than these prices will be considered poor performance. Candidates for this price, might be the closing price or the volume weighted average price (VWAP).

As discussed above and in the book the trades that actually get executed can be different than what was desired. To define what in fact was executed (and at what price) we introduce the variables

- n_1 the portfolio we actually end with where $n_1 \neq v$ normally due to possibly failed limit orders or missing liquidity.
- p are the actual trading prices where trades were executed at. This price will include market impact due to large trading orders.

Based on these variables we will have several constraints. We assume that the initial portfolio value equals the end desired portfolio value at the initial prices or

$$n_0'\pi_0 = v'\pi_0 \,.$$

Since all trading takes place at the prices p and we don't put any cash into the transactions the portfolio we start with (at price p) must equal the one we end with (at price p) or

$$n_0'p = n_1'p\,.$$

With this the *implementation shortfall* is given by

$$\text{implementation shortfall} = v'\pi_1 - n_1'\pi_1 = (v - n_1)'\pi_1\,. \tag{113}$$

We can write the above as

$$\begin{aligned}\text{implementation shortfall} &= (v - n_1)'(\pi_1 - \pi_0) + (v - n_1)'\pi_0 \quad \text{now using } v'\pi_0 = n_0\pi_0 \\ &= (v - n_1)'(\pi_1 - \pi_0) + (n_0 - n_1)'\pi_0 \quad \text{now using } n_0'p = n_1'p \\ &= (v - n_1)'(\pi_1 - \pi_0) + (n_1 - n_0)'(p - \pi_0)\,. \end{aligned} \tag{114}$$

In the case where $v = n_0$ or that the agent already has his desired position then from Equation 113 the implementation shortfall is given by

$$(n_0 - n_1)'\pi_1 = -(n_1 - n_0)'\pi_1 = -(n_1 - n_0)'(\pi_1 - p)\,,$$

since $n_1'p = n_0'p$.

Chapter 15 (Prospective Trading Costs and Execution Strategies)

Notes on the text

Notes on models of order splitting (slicing) and timing

In this section, we assume that the model of midquote (fair price) dynamics, when we trade s_t shares, is given by

$$m_t = m_{t-1} + \mu + \lambda s_t + \varepsilon_t \,. \tag{115}$$

Here μ is the drift of the security. The variable $\lambda > 0$ is the market impact coefficient which affects how the market adjusts the fair price m_t due to our trading. When we submit an order for s_t shares, the trade price is assumed to follow (for the models in this section)

$$p_t = m_t + \gamma s_t \,. \tag{116}$$

At the timestep t we will submit orders to trade s_t shares and desire to trade a total of $\bar{s}$ shares. The constraint between each "order" of s_t and the total desired order $\bar{s}$ is then that $\bar{s} = \sum_{t=1}^{T} s_t$. The problem we attempt to solve is to find an optimal way to split our total order up into pieces under the constraint above. Optimal in this case means that we want to pay the smallest expected cost for our total of $\bar{s}$ shares. That is our problem is to find the order sizes $s_1, s_2, \cdots s_{T-1}, s_T$ to

$$\min_{s_1, s_2, \cdots, s_{T-1}, s_T} E_t \left[\sum_{t=1}^{T} p_t s_t \right] \,. \tag{117}$$

If we assume that s_t can be determined before the period of trading and is predetermined we can pass the expectation into the summation to get

$$\min_{s_1, s_2, \cdots, s_{T-1}, s_T} \sum_{t=1}^{T} s_t E_t[p_t] \,.$$

From our trade price assumption given in Equation 116 we have that $E[p_t] = E[m_t] + \gamma s_t$, thus we now need to evaluate $E[m_t]$. To do this note that we can write m_t using Equation 115 over and over as

$$\begin{aligned}
m_1 &= m_0 + \mu + \lambda s_1 + \varepsilon_1 \quad \text{so} \\
m_2 &= m_1 + \mu + \lambda s_2 + \varepsilon_2 = m_0 + 2\mu + \lambda \sum_{i=1}^{2} s_i + \sum_{i=1}^{2} \varepsilon_i \\
m_3 &= m_0 + 3\mu + \lambda \sum_{i=1}^{3} s_i + \sum_{i=1}^{3} \varepsilon_i \\
&\vdots \\
m_t &= m_0 + t\mu + \lambda \sum_{i=1}^{t} s_i + \sum_{i=1}^{t} \varepsilon_i \,.
\end{aligned}$$

Using this last expression we compute that

$$E[m_t] = m_0 + t\mu + \lambda \sum_{i=1}^{t} s_i \,.$$

Thus our minimization problem becomes

$$\min_{s_1, s_2, \cdots, s_{T-1}, s_T} \sum_{t=1}^{T} s_t \left(m_0 + t\mu + \lambda \sum_{j=1}^{t} s_j + \gamma s_t \right) \,.$$

Note that the first term is

$$\sum_{t=1}^{T} s_t m_0 = m_0 \sum_{t=1}^{T} s_t = m_0 \bar{s} \,,$$

and is independent of the selection of the individual values of s_t and so can be dropped from the optimization. With this realization our optimization problem then becomes

$$\min_{s_1, s_2, \cdots, s_{T-1}, s_T} \sum_{t=1}^{T} s_t \left(t\mu + \lambda \sum_{j=1}^{t} s_j + \gamma s_t \right) \,. \quad (118)$$

Define the objective function in the above optimization problem as $\mathcal{S}$ (for summation). If $T = 3$ then the summation expression above becomes

$$\begin{aligned} \mathcal{S} &= s_1(\mu + \lambda s_1 + \gamma s_1) + s_2(2\mu + \lambda(s_1 + s_2) + \gamma s_2) + s_3(3\mu + \lambda(s_1 + s_2 + s_3) + \gamma s_3) \\ &= \mu(s_1 + 2s_2 + 3s_3) + \gamma(s_1^2 + s_2^2 + s_3^2) + \lambda(s_1^2 + s_2^2 + s_3^2 + s_1 s_2 + s_1 s_3 + s_2 s_3) \,. \end{aligned}$$

If $\mu = 0$ then each of the unknowns s_1, s_2, and s_3 appear in the same manner i.e. no s_j has any stronger influence over the value of the objective function than any other and we expect s_j to be the *same* for each value of j. Note that this is not true if $\mu \neq 0$. For example when $T = 3$ when $\mu \neq 0$ our optimization objective has the term

$$\mu(s_1 + 2s_2 + 3s_3) \,,$$

which gives more weight to s_3 relative to s_1 and s_2 (due to the coefficient of 3). Under the assumption that $s_1 = s_2 = s_3 = \cdots = s_{T-1} = s_T = s$ then to make $\bar{s} = \sum_{t=1}^{T} s_t$ with $s_t = s$ means that our optimal solution is $s_t = \frac{\bar{s}}{T}$.

Exercise 15.1 (variations in the permanent impact parameter)

In the case of the problem describe we need to find s_1, s_2, s_3 to solve the optimization problem given by Equation 117. Under the assumed midquote dynamics of $m_t = m_{t-1} + \lambda_t s_t + \varepsilon_t$ and trade price $p_t = m_t$ this simplifies to

$$\begin{aligned} \min_{s_1,s_2,s_3} E_t \left[\sum_{t=1}^{3} p_t s_t \right] &= \min_{s_1,s_2,s_3} E_t \left[m_1 s_1 + m_2 s_2 + m_3 s_3 \right] \\ &= \min_{s_1,s_2,s_3} E_t \left[(m_0 + \lambda_1 s_1 + \varepsilon_1) s_1 + (m_1 + \lambda_2 s_2 + \varepsilon_2) s_2 + (m_2 + \lambda_3 s_3 + \varepsilon_3) s_3 \right] \\ &= \min_{s_1,s_2,s_3} E_t \left[(m_0 + \lambda_1 s_1 + \varepsilon_1) s_1 + (m_0 + \lambda_1 s_1 + \varepsilon_1 + \lambda_2 s_2 + \varepsilon_2) s_2 + (m_1 + \lambda_2 s_2 + \varepsilon_2 + \lambda_3 s_3 + \varepsilon_3) s_3 \right] \\ &= \min_{s_1,s_2,s_3} E_t \left[(m_0 + \lambda_1 s_1 + \varepsilon_1) s_1 + (m_0 + \lambda_1 s_1 + \varepsilon_1 + \lambda_2 s_2 + \varepsilon_2) s_2 + (m_0 + \lambda_1 s_1 + \varepsilon_1 + \lambda_2 s_2 + \varepsilon_2 + \lambda_3 s_3 + \varepsilon_3) s_3 \right] \\ &= \min_{s_1,s_2,s_3} E_t \left[(m_0 + \lambda_1 s_1 + \varepsilon_1) s_1 + (m_0 + \lambda_1 s_1 + \lambda_2 s_2 + \varepsilon_1 + \varepsilon_2) s_2 + (m_0 + \lambda_1 s_1 + \lambda_2 s_2 + \lambda_3 s_3 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3) s_3 \right] \,. \end{aligned}$$

Since $s_1+s_2+s_3 = 1$ the terms with m_0 don't affect the optimization (as they are independent of s_i) and using the fact that $E_t\varepsilon_i = 0$ for $i = 1, 2, 3$ we find we want to minimize an expression like

$$\lambda_1 s_1^2 + \lambda_1 s_1 s_2 + \lambda_2 s_2^2 + \lambda_1 s_1 s_3 + \lambda_2 s_2 s_3 + \lambda_3 s_3^2 \,,$$

still subject to the constraint that $s_1 + s_2 + s_3 = 1$. To solve this problem we will introduce Lagrange multipliers by first forming our Lagrangian $\mathcal{L}$

$$\mathcal{L} \equiv \lambda_1 s_1^2 + \lambda_1 s_1 s_2 + \lambda_2 s_2^2 + \lambda_1 s_1 s_3 + \lambda_2 s_2 s_3 + \lambda_3 s_3^2 - \lambda(s_1 + s_2 + s_3 - 1) \,.$$

The introduced Lagrangian multiplier parameter, λ, is not related to the market impact parameters λ_i for $i = 1, 2, 3$. Then with this definition of $\mathcal{L}$ to optimize we require solving the partial derivatives of $\mathcal{L}$ with respect to s_1, s_2, s_3, and λ all set equal to zero or the equations

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial s_1} &= 2\lambda_1 s_1 + \lambda_1 s_2 + \lambda_1 s_3 - \lambda = 0 \\
\frac{\partial \mathcal{L}}{\partial s_2} &= \lambda_1 s_1 + 2\lambda_2 s_2 + \lambda_2 s_3 - \lambda = 0 \\
\frac{\partial \mathcal{L}}{\partial s_3} &= \lambda_1 s_1 + \lambda_2 s_2 + 2\lambda_3 s_3 - \lambda = 0 \\
\frac{\partial \mathcal{L}}{\partial \lambda} &= -(s_1 + s_2 + s_3 - 1) = 0 \,.
\end{aligned}$$

Solving these four equations for s_1, s_2, s_3, and λ gives the expressions quoted.