Dynamic Portfolio Management using Optimal Control with Quadratic Costs

Peter Fraenkel, Pragma Securities

1 Introduction

Recent comments by academics and financial practitioners have highlighted [6, 3, 4] the importance of combining portfolio and execution optimizations in order to capture the effects of signal decay, transaction costs and covariant risk of time-dependent holdings. In this note, we show how the mean-variance optimizations traditionally applied separately to portfolios[7] and execution trajectories[1] can be cleanly combined within the framework of optimal control theory.

2 Optimal Control in Linear Systems with Quadratic Cost

In this section, we review an application of dynamic programming and optimal control. For more information, see [2], whose notation we follow closely here, or [8], among others. Consider a generalized system described at time k by a state vector x_k. We allow the system to evolve over time as

$$x_{k+1} = A_k x_k + B_k u_k + w_k \tag{1}$$

for $k \geq 0$ where A_k and B_k are known but possibly time-dependent matrices, w_k is an uncorrelated random increment of mean zero, and u_k is a control vector which we apply in order to minimize an overall cost

$$V = \sum_{k=0}^{N} x_k' Q_k x_k + u_k' R_k u_k \tag{2}$$

that is quadratic in both state x_k and control u_k, with respective coefficient matrices Q_k and R_k, which are deterministic but may vary in time. This model, known as Linear-Quadratic Regulation or LQR is a reasonable description of systems that we wish to control towards some desired state (here $x = 0$) but where there is a cost both to exerting that control and of deviation from the desired state.

To search for an optimal set of controls

$$u_k^* = \arg\min_{\{u_k\}} \mathcal{E}_{\{w_k\}} \left(\sum_{k}^{N} x_k' Q_k k_k + u_k' R u_k \right) \tag{3}$$

using the apparatus of dynamic programming, we solve for u_k^* by working backwards and evaluating at each step the optimal total expected cost $J_k(x_k)$ that, for a given state x_k, will be incurred in steps k through N:

$$\begin{aligned} J_N &= x_N' Q_N x_N \\ J_k(x_k) &= \min_{u_k} \left[x_k' Q x_k + u_k' R_k u_k + \mathcal{E}_{w_k} \left(J_{k+1}(A_k x_k + B_k u_k + w_k) \right) \right] \end{aligned} \tag{4}$$

It turns out that with quadratic costs there is an exact solution to this problem; the optimal control is affine in x:

$$u_k^*(x_k) = -[(B_k' K_{k+1} B_k + R_k)^{-1} B_k' K_{k+1} A_k] x_k \tag{5}$$

where K_k is a symmetric matrix that satisfies the Riccati Equation:

$$K_k = A_k'(K_{k+1} - K_{k+1} B_k (B_k' K_{k+1} B_k + R_k)^{-1} B_k' K_{k+1} A_k + Q_k \tag{6}$$

The total expected cost given this optimal strategy is

$$J_0 = x_0' K_0 x_0 + \sum_{k=0}^{N-1} \mathcal{E}(w_k' K_{k+1} w_k) \tag{7}$$

Note that, while the total cost is affected by the random perturbations w, the optimal control policy depends only on current state x. Our assumption that the w have zero mean causes them to drop out when expectations are taken in equation 4. This result is known as "certainty equivalence."

It will prove useful to have a solution for 6 in the limit of continuous time and constant coefficients. We make the substitutions (see [8]):

$$\begin{aligned} u_k &\to u(t) \\ x_k &\to x(t) \\ x_{k+1} &\to x(t + \delta t) \\ w_k &\to \delta w \\ A &\to (A - I)\delta t \\ B, Q, R &\to B\delta t, Q\delta t, R\delta t \end{aligned} \tag{8}$$

where w is now a Brownian motion. The state obeys the dynamics

$$\dot{x} = Ax + Bu + dw/dt, \tag{9}$$

and the total utility is

$$V = \int_0^T (x'Qx + u'Ru)dt, \tag{10}$$

leading to an optimal control

$$\begin{aligned} u^* &= -R^{-1} B' K x \\ 0 &= \dot{K} + Q + A'K + KA - KBR^{-1}B'K \end{aligned} \tag{11}$$

To develop an intuition for the implications of this mapping, it is useful to look at a steady state solution, where there is no deterministic time dependence, and no fixed time alloted for completion. We have already assumed constant coefficients; we now further assume no breaking of time symmetry through boundary conditions. Additionally, let $A = 0, B = \mathcal{I}$, so $u = \dot{x}$ and (11) becomes

$$Q = KR^{-1}K$$

Recall that K is symmetric; assume that Q is both symmetric and positive definite, and that R is diagonal and positive. Pre- and post-multiply both sides by $R^{-\frac{1}{2}}$:

$$R^{-\frac{1}{2}}QR^{-\frac{1}{2}} = (R^{-\frac{1}{2}}KR^{-\frac{1}{2}})^2$$

The left-hand side can be diagonalized as

$$R^{-\frac{1}{2}}QR^{-\frac{1}{2}} = UM^2U' = (UMU')^2$$

where U is an orthonormal matrix and M^2 is diagonal. It follows that

$$\begin{aligned} K &= R^{\frac{1}{2}}UMU'R^{\frac{1}{2}} \\ u &= -R^{-1}Kx = -R^{-\frac{1}{2}}UMU'R^{\frac{1}{2}}x \end{aligned} \tag{12}$$

or, defining a matrix square-root via the diagonalization above,

$$\dot{x} = -R^{-\frac{1}{2}}\sqrt{R^{-\frac{1}{2}}QR^{-\frac{1}{2}}}R^{\frac{1}{2}} \cdot x \tag{13}$$

3 Application to Optimal Liquidation

Almgren and Chriss[1] formalize optimal liquidation of stock holdings x_k as the minimization of a utility cost, which can be written in the continuous limit as

$$V = \int_0^T (\lambda\sigma^2 x^2 + \eta\dot{x}^2)dt, \tag{14}$$

where λ is a risk aversion, σ is the stock volatility and η is the coefficient for price impact, which is linear in the rate of trading $\dot{x}$. This minimization can be expressed directly in the LQR framework with $A = 0$, $B = 1$, $Q = \lambda\sigma^2$, $R = \eta$ and $w = 0$. Equation 11 reduces to

$$0 = \dot{K} + \lambda\sigma^2 - K^2/\eta, \tag{15}$$

which has solution

$$K = \sqrt{\lambda\sigma^2\eta}\coth\left(\sqrt{\frac{\lambda\sigma^2}{\eta}} \cdot (T - t)\right).$$

When optimally controlled, $\dot{x} = -\frac{1}{\eta}Kx$, which implies

$$x = x_0 \frac{\sinh(\kappa \cdot (T-t))}{\sinh(\kappa T)} \tag{16}$$

with $\kappa = \sqrt{\lambda\sigma^2/\eta}$, exactly as in [1].

Note that the solution for $T \to \infty$, ie. with the same variance driven urgency to trade but no hard time limit, is simpler and more intuitive. Setting $\dot{K} = 0$, equation 15 reduces to $K = \sqrt{\lambda\sigma^2\eta}$, and the optimal trajectory is just

$$\dot{x} = -\kappa x \tag{17}$$
$$x = x_0 e^{-\kappa t} \tag{18}$$

We trade in direct proportion to the number of shares remaining to be executed.

The $T \to \infty$ solution for multiple stocks is just equations (12,13) above with

$$\begin{aligned} Q &= \lambda\Omega \\ R &= \text{diag } \eta \end{aligned}$$

where Ω is a covariance matrix, λ is a coefficient of risk aversion and η is a vector of linear price impact coefficients for each stock.

4 Simultaneous Optimization of Portfolio Holdings and Execution

We write the total cost of a changing portfolio of holdings y_k as

$$V = \sum_k -\alpha_k y_k + y_k' Q_k y_k + u_k' R_k u_k \tag{19}$$

with

$$\begin{aligned} Q_k &= \lambda\Omega_k \\ R_k &= \text{diag } \eta_k \end{aligned}$$

where Ω_k and η_k are now time-dependent, α_k is a time-dependent vector of anticipated stock returns. Holdings are adjusted over time by trading

$$y_{k+1} = y_k + u_k$$

In general, any quantity with a k subscript can vary both deterministically and stochastically, but for illustration purposes, we'll now set Ω and η to be constant, and let α evolve randomly

$$\alpha_{k+1} = \alpha_k + \nu_k,$$

where ν is an uncorrelated random variable with mean zero.

Let $\tilde{y}$ be the instantaneously optimal holdings, given the current value of α:

$$\begin{aligned}\tilde{y}_k &= \arg\min_{y_k} -\alpha_k y_k + y_k' Q y_k \\ &= \frac{1}{2} Q^{-1} \alpha_k \end{aligned} \quad (20)$$

If there were no transaction costs, we would want $y_k = \tilde{y}_k$ at all times. In practice, we will have to accept some displacement $x_k = y_k - \tilde{y}_k$ from instantaneous optimality, with time evolution

$$x_{k+1} = x_k + u_k + w_k \quad (21)$$

The random perturbation $w_k = -\frac{1}{2} Q^{-1} \nu_k$ is still uncorrelated.

Rewriting equation 19 with $y_k = \tilde{y}_k + x_k$, we have

$$U = \sum_k [x_k' Q x_k + u_k' R u_k] + [2\tilde{y}_k Q x_k - \alpha x_k] + [\tilde{y}_k' Q \tilde{y}_k - \alpha \tilde{y}_k] \quad (22)$$

Note that the second set of terms in square brackets is identically zero and that the the third set of terms is a function only of $\tilde{y}$ and thus completely unaffected by our choice of u. We thus get to ignore all but the first terms; the minimization maps onto continuous LQR with $A = 0$, $B = 1$; and we have a trading strategy given by equation 13. In each time period, the ideal portfolio $\tilde{y}$ moves due to fluctuations in α, while our actual holdings y are adjusted by trading towards the *instantaneous* target portfolio $\tilde{y}$ in matrix proportionality to the distance $y - \tilde{y}$. In general, we will never actually arrive at $\tilde{y}$, but our trading strategy represents the optimal extraction of utility given α, risk and trading costs.

Engle and Ferstenberg[3] make use of the cancellation noted in equation 22 above, thereby arriving as we do at an optimization that depends only on displacement from optimality. However, they make the unnecessary assumption that that α and thus $\tilde{y}$ are constant, while also imposing the constraint that all trading complete within an arbitrary period T that is itself much shorter than the eventual holding period. This results in a static trajectory for y towards a constant target. In [4], we argued that these assumptions were unrealistic; here we note that they are unnecessary as long as the random perturbations in $\tilde{y}$ have mean zero.

In practice, the mean zero assumption can be violated quite easily:

- There may be constraints on $\tilde{y}$, in which case it will not be simply proportional to α.
- Reasonable processes for α will be mean reverting and thus not uncorrelated; without mean-reversion, α will be unbounded, which is both financially and mathematically unreasonable.
- Ω and η may be time-varying.

None of these situations prevent us from using dynamic programming to solve for the optimal trading strategy, but in general the solution will involve numerical computation. In particular, as noted in [4], portfolio constraints will necessitate incorporation of full portfolio risk and α directly in the optimization. In the very common case of mean-reverting alpha, LQR does give us some additional useful tools, as discussed in the next section.

5 Mean reversion of distance from optimality

It is reasonable to consider a geometric mean-reverting process for alpha (and thus for $\tilde{y}$)

$$\delta\alpha = -\zeta\alpha\delta t + \sigma_\alpha \alpha \delta z, \tag{23}$$

where δz is a Brownian motion. We may infer $\zeta = \frac{1}{2}\sigma^2$ from an assumption of constant available alpha signal, $\delta(\alpha \cdot \alpha) = 0$.

While the LQR framework does not directly support mean reversion in $\tilde{y}$, it does allow it in $x = y - \tilde{y}$ in the state dynamics of equation 9 via a diagonal matrix A, which for simplicity we earlier assumed to be zero. Instead, we will assume that it is a negative multiple of the identity matrix, $A = -g\mathcal{I}$. If we retain A while following the procedure used to derive equation 13, we obtain

$$K = R^{\frac{1}{2}}\sqrt{R^{-\frac{1}{2}}(Q + A^2R)R^{-\frac{1}{2}}}R^{\frac{1}{2}} + AR \tag{24}$$

$$u = -R^{-1}Kx \tag{25}$$

where the square root is defined through the diagonalization procedure given above. In one dimension,

$$u = -(\sqrt{\lambda\sigma^2/\eta + g^2} - g) \cdot x \tag{26}$$

We will return later to the question of how mean reversion in x can relate to mean reversion in α.

6 Grinold's Dynamic Portfolio Analysis

In the following discussion of Grinold's recent article[8], we stick as closely as possible to his notation. He posits a mean-reverting process for optimal holdings, writing

$$\delta m = -g \cdot m(t - \Delta t) \cdot \Delta t + u(t), \tag{27}$$

where m is a vector of instantaneously optimal holdings, g is a scalar reversion coefficient and $u(t)$ is a random process of mean zero. He expresses the conservation of α as an equilibrium condition for the dollar variances of u and m, $\omega_u^2 = 2g\omega_m^2\Delta t$. He suggests that actual holdings follow the trading policy

$$\Delta p = b \cdot \{m(t) - p(t - \Delta t)\}\Delta t \tag{28}$$

for some optional trading rate constant b (renamed here from Grinold's d for later clarity when using the notation of calculus) and assumes annual trading costs proportional to the variance of the time rate of change of the actual holdings

$$c_p = \frac{\chi}{2}\omega^2_{\Delta P/\Delta t}. \tag{29}$$

Noting that this expression differs from our usual assumption of transaction cost proportional to the integrated square of *trading rate*, rather than the square of portfolio change due both to trading and price fluctuation, we propose a rough equivalence

$$\chi = 2\eta/\sigma^2. \tag{30}$$

Finally, he specifies a risk cost of $\frac{1}{2}\lambda_G\omega_P^2$; we add the G subscript to his risk aversion parameter to make explicit the trivial difference from our convention:

$$\lambda_G = 2\lambda \tag{31}$$

Relating expected portfolio return to the information ratio IR (discussed and defined in [5]), he derives a total utility

$$U_P(b, \omega_M) = \frac{b}{b+g}\left\{IR_M \cdot \omega_M - \frac{\lambda_G/2 + \chi \cdot g \cdot b}{2}\omega_M^2\right\}. \tag{32}$$

At this point, Grinold holds d constant and solves for an optimal value of ω_M,

$$\hat{\omega}_m(b) = \frac{IR_M}{\lambda_G + \chi \cdot g \cdot b}, \tag{33}$$

the optimal rate of trading towards which is given by

$$\Delta p = \hat{b} \cdot \{\hat{m}(t) - p(t - \Delta t)\}\Delta t \tag{34}$$

$$\hat{b} = \sqrt{\lambda_G/\chi} \tag{35}$$

It is our opinion that this step actually violates one of the central assumptions of the paper, namely that m is indeed the model portfolio that one would possess in a ideal world of costless transactions. We would fork off from his derivation after his equation D-5 by keeping the usual relation

$$\omega_m = IR_m/\lambda_G, \tag{36}$$

independent of whatever d we end up using as we try to track m. We now have

$$U_P(b, \omega_M) = \frac{1}{2}\frac{b \cdot \omega_m^2}{b+g}\{\lambda_G - \chi \cdot g \cdot b\} \tag{37}$$

which achieves its maximum at

$$b^* = \sqrt{\lambda_G/\chi + g^2} - g \tag{38}$$

Pleasantly, this is identical to equation 26 with the equations 30 and 31.

7 Mean reversion and optimal trading

It may be, *prima facie*, surprising that identical optimal trading strategies result from assumptions, respectively, of mean-reverting optimal holdings and of mean reverting distance from optimal holdings, with the same proportionality constant in both cases. To see intuitively how this might happen, recall that an Ornstein-Uhlenbeck process r_t, where

$$dm_t = -g \cdot m_t dt + \sigma dW_t, \tag{39}$$

approaches a running exponentially-weighted average of a Brownian motion in the long time limit:

$$m_t = \int_0^\infty \sigma e^{-gs} dW_{t-s} \tag{40}$$

Similarly, a trading strategy for p_t that is proportional to distance from optimality is essentially a running average of the optimal portfolio m_t:

$$p_t = \int_0^\infty e^{-b \cdot s} m_{t-s} ds \tag{41}$$

$$= \int_0^\infty \int_0^\infty e^{-bu} e^{-gv} dW_{t-u-v} du \tag{42}$$

A dynamically optimal portfolio is thus a lowpass-filtered instantaneously optimal portfolio, which is a lowpass-filtered noise process.

References

[1] R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–39, 2001.

[2] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 2005.

[3] Robert F. Engle and Robert Ferstenberg. Execution Risk. *SSRN eLibrary*, 2006.

[4] Peter Fraenkel and Lee Maclin. *Correcting the structural shortcomings of today's portfolio algorithms*, pages 115–122. The Trade, 2007.

[5] R. C. Grinold and R. N. Kahn. *Active Portfolio Management: Quantitative Theory and Applications*. Probus Publishing, Chicago, Ill., 1995.

[6] Richard Grinold. Dynamic portfolio analysis. *The Journal of Portfolio Management*, 34(1):12–26, 2007.

[7] H. M. Markowitz. *Portfolio Selection; efficient diversification of investments*. Wiley, New-York, 1959.

[8] Richard Weber. Optimization and control. Course notes, University of Cambridge, 2007.

