Ralph's Kelly-Ross Investments

Ralph believes managing his investments/projects in accordance with the Kelly-Ross theorem gives him the best chance of meeting his financial goals. The Kelly-Ross theorem combines Ross' recovery theorem with Kelly's mutual information theorem. Ralph understands the strategy places demands on his management skills as it involves no arbitrage (positive state prices) and scalable investment projects that span the state space to accommodate construction of Arrow-Debreu assets.

Ross' recovery theorem. Ross' [2011][1] recovery theorem says that in a complete, pure exchange market setting, linear no arbitrage equilibrium state prices convey a representative investor's state probability assignments. That is, state prices convey Markovian state transition probabilities (and preferences regarding timing of consumption and risk) for a representative investor. This is in the spirit of assigning probabilities based on what we know, in other words, maximum entropy probability assignment. The representative investor's beliefs (and risk preferences) are consistent with equilibrium asset (and state) prices and the representative investor's consumption preferences have at least the same ordering as aggregate consumption (the sum of individual investors' consumption). That is, we can characterize "market" beliefs and preferences as if there exists a representative investor.

The key is the pricing kernel which says the (state) price, p_{ij}, per unit probability, f_{ij}, is equal to a personal discount factor, δ, times the ratio of marginal utilities for consumption in the future state, c_j, to current, c_0, where j refers to the future state.

$$\frac{p_{ij}}{f_{ij}} = \delta \frac{U^{'}(c_j)}{U^{'}(c_0)}$$

In other words, a representative investor with wealth or endowment, W_0, solves for optimal consumption subject to a budget or wealth constraint.

$$\max_{c_0, c_j \geq 0} U(c_0) + \delta \sum_{j=1}^{n} f_{ij} U(c_j)$$
$$s.t. \quad c_0 + \textstyle\sum_{j=1}^{n} p_{ij} c_j \leq W_0$$

The first order conditions for the Lagrangian representation of the above optimization problem yield the pricing kernel.

$$\begin{aligned} \lambda &= U^{'}(c_0) \\ \delta f_{ij} U^{'}(c_j) &= p_{ij} U^{'}(c_0) \end{aligned}$$

For Markovian transition probabilities assigned as $F = \frac{1}{\delta} DPD^{-1}$ where D is a diagonal matrix with elements $U^{'}(c_1), \ldots, U^{'}(c_n)$ and with $U^{'}(c_0) = U^{'}(c_i)$,

[1] Ross, S. 2011, "The recovery theorem," MIT working paper.

then the pricing kernel for the representative investor is[2]

$$\frac{p_{ij}}{f_{ij}} = \delta \frac{U^{'}(c_j)}{U^{'}(c_i)}$$

State transition probability assignment follows from eigensystem decomposition of the dynamic system of state prices P along with the requirement the rows of F sum to one.

$$P\zeta = \delta\zeta$$

where, by the Perron-Frobenius theorem, ζ is the positive-valued eigenvector associated with the largest eigenvalue δ. The Perron-Frobenius theorem says for a nonnegative matrix the largest eigenvalue and its associated eigenvector are nonnegative. Since P is a matrix of state prices, P is a nonnegative matrix (otherwise, there exist arbitrage opportunities).

Let ι be a vector of ones and recall eigenvectors are scale-free, $P(\alpha\zeta) = \delta(\alpha\zeta)$ implies $P\zeta = \delta\zeta$. Then, we can write

$$D^{-1}\iota = \zeta$$

with ζ scaled appropriately. Notice, the pricing kernel is also scale-free as only ratios of marginal utilities enter. Collecting terms, we have

$$\begin{aligned} P\zeta &= \delta\zeta \\ PD^{-1}\iota &= \delta D^{-1}\iota \\ \frac{1}{\delta}DPD^{-1}\iota &= \iota \\ F\iota &= \iota \end{aligned}$$

which confirms that F is a proper probability assignment as the terms are nonnegative and sum to one.

This eigensystem decomposition of P follows directly from the pricing kernel

[2] Constant relative risk aversion is represented as

$$U(c) = \frac{c^{1-r}}{1-r}$$

where $r \to 1$ leads to $U(c) = \ln c$. Constant relative risk aversion is attractive as a change in wealth leads to no change in the relative composition of an individual's portfolio (fraction of wealth invested in various assets). For constant relative risk aversion, relative marginal utility is

$$\frac{U^{'}(c_j)}{U^{'}(c_i)} = \left(\frac{c_j}{c_i}\right)^{-r}$$

and logarithmic relative marginal utility is

$$\frac{U^{'}(c_j)}{U^{'}(c_i)} = \left(\frac{c_j}{c_i}\right)^{-1}$$

and risk preference independence over initial states.

$$\begin{aligned}\frac{p_{ij}}{f_{ij}} &= \delta\frac{U^{'}(c_j)}{U'(c_i)}\\ p_{ij}\frac{U^{'}(c_i)}{U'(c_j)} &= \delta f_{ij}\end{aligned}$$

Since f_{ij} is a probability distribution given initial state i, $\sum_j f_{ij} = 1$. Therefore,

$$\begin{aligned}\sum_j p_{ij}\frac{U^{'}(c_i)}{U'(c_j)} &= \delta\sum_j f_{ij} = \delta\\ p_{i1}\frac{U^{'}(c_i)}{U'(c_1)} + \cdots + p_{in}\frac{U^{'}(c_i)}{U'(c_n)} &= \delta\\ p_{i1}\frac{1}{U'(c_1)} + \cdots + p_{in}\frac{1}{U'(c_n)} &= \delta\frac{1}{U'(c_i)}\end{aligned}$$

For initial states, $i = 1, \ldots, n$, we have n equations. In matrix form, this is

$$P\begin{bmatrix}\frac{1}{U'(c_1)}\\ \vdots\\ \frac{1}{U'(c_n)}\end{bmatrix} = \delta\begin{bmatrix}\frac{1}{U'(c_1)}\\ \vdots\\ \frac{1}{U'(c_n)}\end{bmatrix}$$

This is the eigensystem decomposition of P

$$PD^{-1}\iota = \delta D^{-1}\iota$$

where the eigenvector associated with the largest eigenvalue δ is

$$\begin{aligned}D^{-1}\iota &= \begin{bmatrix}\frac{1}{U'(c_1)} & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & \frac{1}{U'(c_n)}\end{bmatrix}\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}\\ &= \begin{bmatrix}\frac{1}{U'(c_1)}\\ \vdots\\ \frac{1}{U'(c_n)}\end{bmatrix}\end{aligned}$$

Kelly's long-run wealth maximizing investment strategy. A Kelly investment strategy produces maximum long-run wealth where weights, k, on Arrow-Debreu portfolios match assigned state probabilities, $k = p$, and information revising state probabilities, $p(s)$, is met with portfolio rebalancing accordingly.

Long-run wealth maximization implies maximizing the geometric mean of portfolio returns or expected compound returns

$$\begin{array}{ll} \max\limits_{k} & \prod\limits_{i-1}^{n} \left(\frac{k_i}{y_i}\right)^{p(s_i)} \\ s.t. & \sum_{i=1}^{n} k_i = 1 \end{array}$$

or equivalently, and perhaps more familiarly, (arithmetic) mean of logarithmic returns

$$\begin{array}{ll} \max\limits_{k} & E\left[\log \frac{k}{y}\right] \equiv p(s)^T \log(\Omega k) \\ s.t. & \sum_{i=1}^{n} k_i = 1 \end{array}$$

where k_i is the fraction of wealth invested in project i, $y > 0$ is a vector of no-arbitrage state prices (or Arrow-Debreu prices) derived from $Ay = x$, Ω is a diagonal matrix comprised of $\frac{1}{y_i}$, A is an $n \times n$ matrix of returns with rows referring to projects and columns to states, and x is a vector of investments (normalized to unity).[3]

The first order conditions for the Lagrangian associated with the logarithmic returns frame above regarding long-run wealth maximization is

$$\mathcal{L} = \sum_{i=1}^{n} p_i \ln\left(k_i \frac{1}{y_i}\right) - \lambda\left(\sum_{i=1}^{n} k - 1\right)$$

are

$$\frac{p_i}{k_i} - \lambda = 0, \quad \text{for all } i$$

Since $\sum k_i = 1 = \sum \frac{p_i}{\lambda} = \frac{1}{\lambda}$, $\lambda = 1$ and $k_i = p_i$. In other words, probability assignment to state i identifies the optimal fractional investment in state i.

[3] Weights on the nominal assets are

$$w^T = k^T \Omega A^{-1}$$

where maximizing compound (or geometric mean) return on investment is

$$\begin{array}{ll} \max\limits_{w} & G[r] = \prod\limits_{j=1}^{n} \left(w^T A_j\right)^{p_j} \\ s.t. & w^T \iota = 1 \end{array}$$

or equivalently maximizing the arithmetic mean (expected value) of the natural logarithm of returns is

$$\begin{array}{ll} \max\limits_{w} & E[r] = \sum_{j=1}^{n} p_j \ln\left(w^T A_j\right) \\ s.t. & w^T \iota = 1 \end{array}$$

and w_i (element i of vector w) is portion of wealth invested in project i (this quantity may be negative which translates into borrowing against its future payoff, that is borrow the investment amount and return the payoff to the lender), A_j is column j of A or return (payoff on investment equal to one) on project i in state j (so that $\sum_{i=1}^{m} w_i r_{ij} = w^T A_j$ is the return on the portfolio of projects in state j), p_j is the probability assigned to state j, and $G[r] = \exp(E[r])$.

Mutual information theorem. The mutual information theorem says the expected gain from information equals mutual information.

$$E[r \mid z] - E[r] = I(s; z)$$

where expected long-run wealth is maximized by equating portfolio weights on Arrow-Debreu assets, k_j, with assigned state probabilities, $\Pr(s_j)$, and state transition probabilities, $\Pr(s_j \mid z_i)$.[4]

$$\begin{aligned} E[r_i] &= \sum_{j=1}^{n} \Pr(s_j \mid z_i) \log \frac{k_j}{y_{ij}} \\ &= \sum_{j=1}^{n} \Pr(s_j \mid z_i) \log \frac{\Pr(s_j)}{y_{ij}} \end{aligned}$$

for initial state i and transition (from state i to j) state price y_{ij} (returns are $\frac{1}{y_{ij}}$ as the payoff is one and a state price of y_{ij}). Unconditional expected long-run returns are

$$E[r] = \sum_{i=1}^{n} \Pr(z_i) E[r_i]$$

On the other hand, information z fully accounts for state transition likelihoods, $\Pr(s_j \mid z_i)$, and results in

$$E[r_i \mid z_i] = \sum_{j=1}^{n} \Pr(s_j \mid z_i) \log \frac{\Pr(s_j \mid z_i)}{y_{ij}}$$

and

$$E[r \mid z] = \sum_{i=1}^{n} \Pr(z_i) E[r_i \mid z_i]$$

The expected gain from information z is

$$E[r \mid z] - E[r]$$

Mutual information is

$$I(s; z) = H(s) + H(z) - H(s, z)$$

where $H(\cdot) = -\sum_{j=1}^{n} \Pr(\cdot) \log \Pr(\cdot)$ or entropy associated with states (s), information (z), or states and information jointly (s, z). While expected gains and mutual information appear to have entirely different units, a Kelly investment strategy produces an equivalence between the two seemingly disparate quantities.

[4] $\Pr(s_j \mid z_i)$ is equivalent to F_{ij} in the discussion of Ross' recovery theorem. The notation adopted here is intended to emphasize the conditioning on knowledge of the current state as well as price information dynamics.

The Kelly-Ross theorem. The Kelly-Ross theorem says if state transition probabilities are assigned in accordance with Ross' recovery theorem to reflect price dynamic information (z)

$$F = \frac{1}{\delta} D P D^{-1}$$

where spanning is satisfied and $P > 0$, and investments are made in accordance with the Kelly criterion to maximize expected long-run value based on known initial states then the expected long-run rate of return, $E[r \mid z]$, equals $\log \frac{1}{\delta}$ and the expected long-run growth rate, $\exp(E[r \mid z])$, equals $\frac{1}{\delta}$.

Notice, this says if Ralph assigns probabilities that take into account the information in prices and employs a Kelly investment strategy (to maximize expected long-run wealth) then the long-run expected growth rate is immediately identified as $\frac{1}{\delta}$, the reciprocal of the largest eigenvalue of P, the transition state price matrix. Hence, remarkably, as in the mutual information theorem, returns and probabilities are linked even though they seem to involve different units.

Scenario 1. Suppose Ralph believes returns associated with initial state one are

$$A_1 = \begin{bmatrix} 1.02 & 1.02 \\ 1.1 & \frac{1}{1.1} \end{bmatrix}$$

and for initial state two returns are

$$A_2 = \begin{bmatrix} 1.02 & 1.02 \\ 1 & 1.1 \end{bmatrix}$$

Throughout investments are normalized to unity, that is,

$$x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Suggested: 1. Solve $Ay = x$ for each initial state and create the transition state price matrix, $P = \begin{bmatrix} y_1^T \\ y_2^T \end{bmatrix}$.

2. Find the largest eigenvalue, δ, and associated eigenvector, ζ, for P.

3. Assign state transition probabilities based on state price information.

$$F = \frac{1}{\delta} D P D^{-1}$$

where D^{-1} is a diagonal matrix comprised of the elements of ζ.

4. Employ a Kelly investment strategy given the initial state is known and find $E[r_i] = F_i \log\left(\Omega_i F_i^T\right)$ for each initial state i where Ω_i is a diagonal matrix with elements $\frac{1}{y_i}$ (the reciprocal of state prices associated with initial state i) and F_i is the ith row of F.

5. Determine $E[r]$ by first determining the unconditional steady-state probabilities, p, then utilize p as weights on $E[r_i]$.

$$\begin{aligned} E[r] &= \sum_{i=1}^{n} p_i E[r_i] \\ &= p^T \begin{bmatrix} E[r_1] \\ \vdots \\ E[r_n] \end{bmatrix} \end{aligned}$$

Hint: p is an eigenvector of F^T associated with an eigenvalue equal to one. That is, $p^T F = p^T$ where p is scaled to add to one, $p^T \iota = 1$.

6. Compare $E[r]$ with $\log \frac{1}{\delta}$ and $\exp(E[r])$ with $\frac{1}{\delta}$.

7. Compare $E[r]$ based on price information with expected log-returns where Ralph's investment strategy utilizes p instead of F_i for each initial state i

$$E[r_i^{pss}] = F_i \log(\Omega_i p)$$

and

$$E[r^{pss}] = p^T \begin{bmatrix} E[r_1^{pss}] \\ \vdots \\ E[r_n^{pss}] \end{bmatrix}$$

How does this relate to mutual information? Hint: Joint probabilities are $p_i F_{ij}$ for $i, j = 1, \ldots, n$.

8. Since the recovery theorem is Markovian, the recovery theorem is reversible. In other words, we can assign transition probabilities in the reverse direction (from current to previous state). Let G denote the reverse transition probability assignment.

$$G = \frac{1}{\delta} D_G P^T D_G^{-1}$$

where D_G^{-1} is a diagonal matrix comprised of the elements of the eigenvector for P^T, say ζ_G, associated with δ. Assign G and verify $E_G[r] = \log \frac{1}{\delta}$ where

$$E_G[r_i] = G_i \log\left(\Omega_i G_i^T\right)$$

and

$$E_G[r] = p^T \begin{bmatrix} E_G[r_1] \\ \vdots \\ E_G[r_n] \end{bmatrix}$$

(in other words, the Kelly-Ross theorem applies in reverse as well as forward).

9. F and G represent a full set of conditional probability distributions from which the joint distribution can be derived. Let P_{ij} denote the joint likelihood associated with $z = i$ and $s = j$. The ratio of joint likelihoods can be expressed in terms of the full set of conditional likelihoods.

$$\frac{P_{ij}}{P_{k\ell}} = \frac{F_{ij}}{F_{i\ell}} \frac{G_{\ell i}}{G_{\ell k}} \equiv c_{k\ell}$$

or

$$\begin{aligned} P_{ij} - \frac{F_{ij}}{F_{i\ell}} \frac{G_{\ell i}}{G_{\ell k}} P_{k\ell} &= 0 \\ P_{ij} - c_{k\ell} P_{k\ell} &= 0 \end{aligned}$$

Set $i = j = 1$ and create a system of equations with all permutations of k and ℓ other than $k = \ell = 1$ along with one equation indicating the sum of P_{ij} equals one. Now, solve

$$AP = b$$

where

$$A = \begin{bmatrix} 1 & -c_{12} & 0 & 0 \\ 1 & 0 & -c_{21} & 0 \\ 1 & 0 & 0 & -c22 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

and

$$b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

How does the joint distribution P compare with that in 7?

Scenario 2. Suppose Ralph believes returns associated with initial state one are

$$A_1 = \begin{bmatrix} 1.01 & 1.01 \\ 1.1 & \frac{1}{1.1} \end{bmatrix}$$

and for initial state two returns are

$$A_2 = \begin{bmatrix} 1.05 & 1.05 \\ 1 & 1.1 \end{bmatrix}$$

Suggested: Repeat 1-9 for scenario 2 and compare with scenario 1.

Scenario 3. Suppose Ralph believes returns associated with initial state one are

$$A_1 = \begin{bmatrix} 1.02 & 1.02 & 1.02 \\ 1.1 & \frac{1}{1.1} & 1 \\ 1 & 1.1 & \frac{1}{1.1} \end{bmatrix}$$

for initial state two returns are

$$A_2 = \begin{bmatrix} 1.02 & 1.02 & 1.02 \\ 1 & 1.1 & \frac{1}{1.1} \\ 1.2 & \frac{1}{1.2} & 1 \end{bmatrix}$$

and for initial state three returns are

$$A_3 = \begin{bmatrix} 1.02 & 1.02 & 1.02 \\ 1 & 1.3 & \frac{1}{1.3} \\ 1.2 & \frac{1}{1.2} & 1 \end{bmatrix}$$

Suggested: Repeat 1-9 for scenario 3 and compare with scenarios 1 and 2. (note: the A matrix in 9 is 9×9 rather than 4×4.)

Scenario 4. Suppose Ralph believes returns associated with initial state one are

$$A_1 = \begin{bmatrix} 1.01 & 1.01 & 1.01 \\ 1.1 & \frac{1}{1.1} & 1 \\ 1 & 1.1 & \frac{1}{1.1} \end{bmatrix}$$

for initial state two returns are

$$A_2 = \begin{bmatrix} 1.03 & 1.03 & 1.03 \\ 1 & 1.1 & \frac{1}{1.1} \\ 1.2 & \frac{1}{1.2} & 1 \end{bmatrix}$$

and for initial state three returns are

$$A_3 = \begin{bmatrix} 1.05 & 1.05 & 1.05 \\ 1 & 1.3 & \frac{1}{1.3} \\ 1.2 & \frac{1}{1.2} & 1 \end{bmatrix}$$

Suggested: Repeat 1-9 for scenario 4 and compare with scenarios 1, 2 and 3.

Hint: The largest eigenvalue of P can be determined via an iterative (power) algorithm. Let

$$b_k = P b_{k-1}$$

where, for instance, the initial vector is $b_1 = \iota$. Next, rescale b_k by $\mu_k \equiv \sqrt{b_{k-1}^T P^T P b_{k-1}}$ (recall, eigenvectors are scale-free).

$$\zeta_k = \frac{b_k}{\mu_k}$$

As $k \to n$ the quantity μ_k converges to δ and ζ_k is the associated eigenvector, that is, $P\zeta_k - \mu_k \zeta_k \to 0$.

Alternatively, let $\mu_k \equiv \frac{b_k^T P b_k}{b_k^T b_k}$. This follows as eigensystems are defined by

$$P\zeta_k = \delta \zeta_k$$

Now, multiply both sides by ζ_k^T to generate a quadratic form (scalars on both sides of the equation).

$$\zeta_k^T P \zeta_k = \delta \zeta_k^T \zeta_k$$

Then, isolate the eigenvalue, δ, by dividing both sides by the right-hand side scalar, $\zeta_k^T \zeta_k$, to produce the result. As $k \to n$,

$$\mu_k \equiv \frac{\zeta_k^T P \zeta_k}{\zeta_k^T \zeta_k} \to \delta$$

Scenario 5. Suppose Ralph believes returns associated with initial state one are

$$A_1 = \begin{bmatrix} 1.04 & 1.03 & 1.0244 \\ 1.1 & \frac{1}{1.1} & 1 \\ \frac{1}{1.1} & 1 & 1.1 \end{bmatrix}$$

for initial state two returns are

$$A_2 = \begin{bmatrix} 1.04 & 1.03 & 1.01912 \\ 1.2 & \frac{1}{1.1} & 1 \\ \frac{1}{1.1} & 1 & 1.2 \end{bmatrix}$$

and for initial state three returns are

$$A_3 = \begin{bmatrix} 1.04 & 1.03 & 1.01523 \\ 1.3 & \frac{1}{1.1} & 1 \\ \frac{1}{1.1} & 1 & 1.3 \end{bmatrix}$$

Suggested: Repeat 1-9 for scenario 5 and compare with scenarios 1, 2, 3, and 4.