Kelly Criterion Position Sizing == Mean-Variance Optimization

Max Dama 7/4/09

There are a couple different approaches to determining the best leverage to use for an overall portfolio. In finance 101 they teach Markowitz's mean-variance optimization where the efficient portfolios are along an optimal frontier and the best one among those is at the point where a line drawn from the risk free porfolio/asset is tangent to the frontier. In the 1950s Kelly derived a new optimal leverage criteria inspired by information theory (which had been established by Shannon a few years earlier). The criteria being optimized in these two cases is typically called an 'objective function'- a function of possible asset weights/allocations that outputs a number giving the relative estimated quality of the portfolio weights. However, the two objective functions look quite different (peek ahead if you're unfamiliar). In this note I show the two are approximately equivalent, with the approximation being very close in realistic risk-return scenarios.

I'm ignoring the various other naive/heuristic position sizing approaches which float around the tradersphere- 'half-Kelly', risk-'multiple' based approaches, 130/30 strategies, beginners' basic 100% allocation, etc.

The following are synonyms in the literature.
Mean variance == Modern portfolio theory == Markowitz
Kelly criterion == log optimal == expected geometric growth optimal

Markowitz

Mean variance portfolio optimization maximizes the objective function:

$$\begin{aligned} objective(w) &= E[w*r] - \lambda\,\mathrm{Var}[w*r] \\ &= E[w*r] - \lambda(E[(w*r)^2] - E[w*r]^2) && \text{Computational formula for variance} \\ &\approx E[w*r] - \lambda E[(w*r)^2] && \text{Justification below} \end{aligned}$$

Where λ represents risk aversion- supposedly a known constant like 1. We are allowed to do the last step because in the financial setting, $E[w*r]$ is typically less than 1 i.e. returns are expected to be less than 100%. This causes $E[w*r]^2 << E[w*r]$ so we can ignore it as a round-off error. The fact that $w*r$ is so much less than 1 (since asset weights sum to 1 and returns are less than 1) will be useful later too.

Kelly

This Taylor series formula from Wikipedia will be useful below when we work with Kelly's objective:

$$\log(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} \text{ for } |x| \leq 1, x \neq -1$$

Taylor series can be used to approximate functions such as $\ln(1+x)$ by calculating the first couple terms i.e. $n = \{1, ..., m\}, m < \infty$. They were discovered in the 1300s in India but took another 300 years for European 'mathematicians' to figure out.

Moving on, in contrast to Markowitz above, Kelly maximizes the log growth rate:

$$
\begin{aligned}
objective(w) &= E[\log(1 + w * r)] \\
&= E[\sum_{n=1}^{\infty}(-1)^{n+1}\frac{(w*r)^n}{n}] && \text{As above, assume } w*r < 1 \\
&\approx E[\sum_{n=1}^{2}(-1)^{n+1}\frac{(w*r)^n}{n}] && \text{Throwing away terms corresponding to } n > 2 \\
&= E[(-1)^{1+1}\frac{(w*r)^1}{1} + (-1)^{2+1}\frac{(w*r)^2}{2}] && \text{Expanding the sum} \\
&= E[\frac{(w*r)^1}{1} + (-1)\frac{(w*r)^2}{2}] && \text{Evaluating the two terms } (-1)^k \\
&= E[w*r] - \frac{1}{2}E[(w*r)^2] && \text{Linearity of expectation}
\end{aligned}
$$

This is the same as the final result of Markowitz above except here $\lambda = \frac{1}{2}$.

Where 'Approximately' Breaks Down

First of all the user must obviously have a 'risk aversion' which seeks only to maximize wealth ($\lambda = \frac{1}{2}$).

Now let's look at the places where approximately equals $\approx$ was used in each derivation.

Throwing away $-E[w * r]^2$ in Markowitz will be violated if the expected return is near 100% or higher. Since we're looking at daily returns $E[w * r]$ is almost certainly $< 1\% \implies E[w * r]^2 < .01\%$ so we can basically accept this one as a reasonable approximation.

The looser approximation was in the derivation of Kelly where we threw away terms corresponding to $n > 2$ in the summation. If your strategy has high alpha ($E[w * r] > 0$) and is positively skewed (i.e. fat right tail returns $E[(w*r)^3] > 0$) then this is a bad approximation. This is the case if, let's say, you're Peter Lynch or Warren Buffett and every now and then you pick a 'ten-bagger' while even your unsuccessful picks don't lose much or stay flat. Alternatively you could just have an option strategy which loses money most of the time but every now and then makes a huge win, for example.

Keep in mind that when you use Markowitz instead of Kelly you lose this sensitivity to skewness (and higher moments).