1. Grinold and Kahn(GK)

推导主动管理基本定律的思路为，假设有 N 只股票，M 个信号，用这 M 个信号预测 N 只股票的收益率。记 N 只股票的收益率为 $\mathbf{r}=\left(r_1,\ldots,r_N\right)'$，记这 M 个信号为 $\mathbf{x}=\left(x_1,\ldots,x_M\right)'$，所谓预测，是指得到 $E\left(\mathbf{r}|\mathbf{x}\right)$ 和 $V\left(\mathbf{r}|\mathbf{x}\right)$，即收益率关于信号的条件均值和条件方差。已知 $E\left(\mathbf{r}|\mathbf{x}\right)$，$V\left(\mathbf{r}|\mathbf{x}\right)$ 和风险厌恶系数 λ 的情况下，进行均值方差最优化，得到均值方差最优组合的权重 $\mathbf{w}\left(\mathbf{x}\right)=\dfrac{1}{\lambda}V\left(\mathbf{r}|\mathbf{x}\right)^{-1}E\left(\mathbf{r}|\mathbf{x}\right)$（1），然后计算信息率 $IR\left(\mathbf{x}\right)=\dfrac{\mathbf{w}'E\left(\mathbf{r}|\mathbf{x}\right)}{\sqrt{\mathbf{w}'V\left(\mathbf{r}|\mathbf{x}\right)\mathbf{w}}}=\sqrt{E\left(\mathbf{r}|\mathbf{x}\right)'V\left(\mathbf{r}|\mathbf{x}\right)^{-1}E\left(\mathbf{r}|\mathbf{x}\right)}$（2）。

GK 然后对 $IR\left(\mathbf{x}\right)$ 的平方求无条件期望，把 $\mathbf{x}$ 消掉，中间做了一些假设（假设 $\mathbf{x}$ 的元素均值为 0 方差为 1，各元素之间不相关；假设 $\mathbf{r}$ 的元素均值为 0 方差为 1，各元素之间不相关；假设 $\rho(x_i,r_k)=\mathrm{cov}(x_j,r_k),i,j=1,..,M,k=1,\ldots,N$），得到 $IR^2=IC^2\cdot M$（3），其中，$IC^2=\sum_{k=1}^{N}\rho(x_i,r_k)^2$。

2. Zhuanxin Ding and Martin(DM)

DM 对 GK 做了改进。与其说是改进，不如说是特殊化。特殊化的地方有 2 点。第 1 点，加上时间下标 t，$\mathbf{r}_t$ 和 $\mathbf{x}_t$，相应的，（2）式变为

$$IR_t\left(\mathbf{x}_t\right)=\frac{\mathbf{w}_t'E\left(\mathbf{r}_t|\mathbf{x}_t\right)}{\sqrt{\mathbf{w}_t'V\left(\mathbf{r}_t|\mathbf{x}_t\right)\mathbf{w}_t}}=\sqrt{E\left(\mathbf{r}_t|\mathbf{x}_t\right)'V\left(\mathbf{r}_t|\mathbf{x}_t\right)^{-1}E\left(\mathbf{r}_t|\mathbf{x}_t\right)}\text{（4）。}$$

第 2 点，假设收益率符合多因子模型

$r_{it}=x_{i1t}f_{1t}+x_{i2t}f_{2t}+\ldots+x_{iMt}f_{Mt}+\varepsilon_{it},i=1,\ldots,N,t=1,\ldots,T$（5）。

再做一些假设，$f=E\left(f_{1t},\ldots,f_{Mt}\right)'=E(\left(f_{1t},\ldots,f_{Mt}\right)'|X_t)$（6）

则可以得到

$E\left(\mathbf{r}_t|\mathbf{x}_t\right)=\mathbf{X}_t f$，（7）

$V\left(\mathbf{r}_t|\mathbf{x}_t\right)=\mathbf{X}_t V_f \mathbf{X}_t'+V_\varepsilon$，（8）

将（7）和（8）代入（4）式，再做一些标准化方面的假设（$\mathbf{x}_t$ 的元素均值为 0，方差为 1，V_ε 的对角元素相同等），对（4）式的平方求无条件期望，消掉 $\mathbf{x}_t$，得到，

$IR^2 = f'(\mathbf{I}\sigma^2 / N + V_f)^{-1} f$ （9），其中σ^2为V_ε的对角线元素。

分析（9）式，当$V_f = \mathbf{0}$，即因子收益率不是随机变量而是常数的时候，（9）变为（3）式。当N很大的时候，（9）式的结果与有M个因子收益率$f_{kt}, k = 1, ..., M$，对这M个因子收益率进行均值方差最优化的结果一样。

该文章算是目前见到的对于多因子模型和主动管理定理推导最严谨的，上面只是列了大概内容。列出该文章除了要看看它对主动管理定理的推导之外，另一个重要原因是（7）式和（8）式。（7）式算是给我们目前的做法给了确认。但是仔细一想，（7）式的f的因子收益率的无条件均值，我们用过去36个月的因子收益率来估计估计f，本来可以使用全部历史的因子收益率的估计该无条件均值。我们的做法相当于用 36个月的均值这个信息来预测下期的因子收益率，严格来讲，这是条件均值，从而这是一种因子择时。另外要注意的是上述推导使用了（8）式，即因子收益率的方差矩阵，而我们目前的做法没有考虑阿尔法因子的方差矩阵，这方面可以改进。

3. Jose Menchero(2017)

注意第 1 部分的（2）式，GK 和 DM 从（2）开始，都对$IR(\mathbf{x})$的平方求无条件期望，把$\mathbf{x}$消掉。但是这么做没有必要，（2）本身就包含了主动管理定律想表达的内容。为了看清（2）式的含义，假设$V(\mathbf{r}|\mathbf{x})$是对角矩阵，则（2）式变为

$IR(\mathbf{x}) = \sqrt{\sum_{i=1}^{N} \frac{E(r_i|\mathbf{x})^2}{V(r_i|\mathbf{x})}} = \sqrt{N} \cdot \sqrt{\frac{1}{N}\sum_{i=1}^{N} (\frac{E(r_i|\mathbf{x})}{sde(r_i|\mathbf{x})})^2}$（10），其中$V(r_i|\mathbf{x}) = sde(r_i|\mathbf{x})^2$。

（10）式中，$\sqrt{\frac{1}{N}\sum_{i=1}^{N} (\frac{E(r_i|\mathbf{x})}{sde(r_i|\mathbf{x})})^2}$不正好表示信息$\mathbf{x}$的质量吗？Menchero 将$\sqrt{\frac{1}{N}\sum_{i=1}^{N} (\frac{E(r_i|\mathbf{x})}{sde(r_i|\mathbf{x})})^2}$定义为信息质量$Q$，将$\frac{IR(\mathbf{x})}{Q}$定义为分散度$D$，这一点从（10）可以看出来，在$V(\mathbf{r}|\mathbf{x})$是对角矩阵为对角矩阵的时候，分散度为$\sqrt{N}$。如果不是对角矩阵，则分散度会不一样，可想而知，这是由非对角线元素，即不同股票收益率之间有相关系导致的。所以$IR = D \cdot Q$，Menchero 把他成为新的主动管理基本定律。

4. 回测

每期我们得到的是条件在当期信息$\mathbf{x}_t$下的最优组合的收益率r_{Pt}，r_{Pt}可以表示成

$r_{Pt} = g(\mathbf{x}_t) + u_t$，$E(u_t) = 0, Var(u_t) = \sigma^2$，$u_t$跟$\mathbf{x}_t$独立，其中，$g(\mathbf{x}_t)$为$r_{Pt}$条件在$\mathbf{x}_t$上的均值。我们回测，求的是或者想求的是$r_{Pt}$的均值和方差，以及它们的比值。

这里就已经隐含了r_{Pt}是稳定的时间序列，否则对该时间序列求均值和方差是有问题

的。可以看到

$E(r_{Pt})=E(g(\mathbf{x}_t)),Var(r_{Pt})=Var(g(\mathbf{x}_t))+\sigma^2$（11）。

举例：

（1）某个 CTA 策略，如果我们的信号是如果当日收益率大于前 2 日收益率均值，则第二日买入该股票，持有 1 天，然后平仓。在该情况下，$\mathbf{x}_t$ 只能取 1 个值，该值表示当日收益率大于前 2 日收益率均值，所以 $g(\mathbf{x}_t)$ 为一个常数，$E(r_{Pt})=\sigma^2$。

（2）某个 CTA 策略，如果当日收益率大于前 2 日收益率均值，则第二日买入该股票，持有 1 天，然后平仓；如果当日收益率小于前 2 日收益率均值，则第二日买出该股票，持有 1 天，然后平仓。该情况下，$\mathbf{x}_t$ 能取 2 个值，1 个值表示当日收益率大于前 2 日收益率均值，另一个值表示当日收益率小于前 2 日收益率均值。只要这 2 中情况下收益率均值不一样，$g(\mathbf{x}_t)$ 就不是一个常数，$Var(g(\mathbf{x}_t))$ 也不为 0。

（3）多因子选股策略。同（2），只不过 $\mathbf{x}_t$ 能取很多连续值。

参考文献：

Grinold,Kahn,active portfolio management, Chaper 6,appendix ,2000,2nd editon

Zhuanxin Ding, R. Douglas Martin, The Fundamental Law of Active Management: Redux,2017

Jose Menchero,rethinking the fundamental law of activemanagement,2017