

协方差矩阵的估计与管制技术

蓝海

协方差矩阵的估计，尤其是高维度协方差矩阵的估计，依旧是统计学领域内具备挑战性的开放问题。一个 $p \times p$ 维度的协方差矩阵具备 $\frac{p(p+1)}{2}$ 相互制约的参数。在当前的应用环境中，用以估计协方差矩阵的原始数据的采样数 n ，可能会有 $n < p$ 的情况。这种情况下使用常见的采样平均的方法计算每一个参数是有问题的，尤其是当协方差矩阵的逆矩阵需要求解的情况。比如分类问题，多元线性回归，投资组合以及图形计算中，我们都会遇到这样的问题。正则化技术 (regularization) 是常见的方法。其基本逻辑是加入额外的损失函数或者约束使得计算结果偏离病态、增加鲁邦性的方法。在此，我们介绍几类常见的协方差矩阵的估计与正则化技术。

基于线性多因子模型的协方差矩阵估计

假设有 p 个资产其对应回报率为 $(r^{(1)}, r^{(2)}, \dots, r^{(p)})$