广义矩估计

矩估计法可以通过多种方式推广。
例如，我们可以使用类型为 $n^{-1} \sum_{i=1}^{n} g\left(X_i\right)$ 的平均值来代替样本矩 $n^{-1} \sum_{i=1}^{n} X_i^j$，其中 g 是适当选择的函数。我们也可以使用更一般的 X 的函数来代替平均值。

总之其核心思想是通过解形如 $g(X)=e(\theta)$ 的方程组来确定估计量，其中 $e(\theta)=\mathrm{E}_\theta g(X)$ 是对参数 θ 的期望。

如果参数是 k 维的，那么使用 k 个方程来定义矩估计量似乎是合乎自然的。
问题在于选择哪些函数?
实际上，矩估计法首先将观测数据简化为 k 个函数的值，基于这些简化数据得到矩估计量。如果原始数据不能从这 k 个值中重构，那么这种简化会导致信息的丢失。因此，选择使用哪些函数对于所得估计量的效率至关重要。

一种避免信息丢失的方法是使用比未知参数更多的矩。由于这会导致方程数量多于未知数，因此通常无法找到一个参数值使得样本矩与理论矩完全相等。

相反，我们可以最小化这两类矩之间的距离，例如如下形式的表达式：

$$\sum_{j=1}^{l}\left(\frac{1}{n} \sum_{i=1}^{n} g_j\left(X_i\right)-\mathrm{E}_\theta g_j\left(X_1\right)\right)^2$$

其中 $g_1, ..., g_l$ 是已知的固定函数。估计量 $\hat{\theta}$ 是使该表达式最小化的 θ 值。这种方法被称为广义矩估计法，特别是在计量经济学中被广泛应用。