经济学中巧妙的对数应用

文：James Hamilton　译：陈宗昌

原文链接: *Use of logarithms in economics* by *James Hamilton*

可能你经常会好奇为什么经济学家一直执着于对对数的使用。如果你愿意花一点时间看一些公式和图表的话，你会在下面的段落中找到答案的。

开始，咱们先用一点时间快速复习一下复利率 (compound interest rates) 的概念。假如你投资了一个项目，投资本金 (principal) 是 P_1，期限为一年，年利率为 r%，那么在年终你总共就会有 $P_2 = (1 + r)P_1$ 这个数量的总资金。如果你的年利率为 4%，那么 $P_2 = 1.04P1$。有时候你的利率可能是复合的 (compounded)，例如你在头六个月结束后得到你所有的本金和 2% 的利息，然后在年终结算时，你可以再获得数量为当前账户内总金额 (本金加上前六个月的利息总和) 的 %2 的第二部分利息，这时候你持有的总金额 $P_2 = (1+r/2)(1+r/2)P_1 = (1 + r/2)^2P_1$。若你的利率 $r = 4\%$，那像上述那种复合方式一年复合两次可以让你的年回报率实际增长至 4.04%。如果复合四次的话 $P_2 = (1+r/4)^4P_1$，你的年总利息则可以达到本金的 4.06%。事实证明随着复合频率 (frequency of compounding) n 逐渐成长为一个任意大的值，公式 $(1 + r/n)^n$ 会收敛至一个特定的函数

$$\lim_{n\to\infty}(1 + r/n)^n = e^r \tag{1}$$

这里的 e(约等于为 2.72) 是一个极其特别的数字 (通常被称为自然常数或欧拉常数)，它自身拥有包括此处 (1) 在内的很多神奇而又美丽的性质。所以，如果你的利率连续复合，那么在年终你账户里的总金额会增长 $P_2/P_1 = e^r$ 倍。依然用 4% 来举例，如果采用连续复利 (continuous compounding) 的方式，年终你的实际利率会为 4.08%，这比四次复利操作所获得的 4.06% 的回报率 (rate of return) 和无复利操作的 4.0% 的回报率都要高一些。

取自然对数仅仅是对上述操作的一种逆运算：$\ln(P_2/P_1)$，或者说由于取对数后分数的比值结果即为分子与分母分别取对数后双方的差值，所以这里 $\ln(P_2) - \ln(P_1) = r$。再换句话说，对第二年和第一年的账户金额分别取对数并求出它们差值的操作，刚刚好是在计算这一年间的的投资回报率。当然这些讨论都是在连续复利的场景下进行的。

对于一个很小的 r 来说，连续复利的回报与不进行连续复利操作的回报几乎相等，所以他们的对数差几乎和百分比变化率相同。回到之前的那个例子，该示例中的百分比变化率为 $(P_2 - P_1)/P_1 = 0.0408$，对数变化为 $\ln(P_2) - \ln(P_1) = 0.04$。所以每当你看到一个量度为对数的图表时，按其图中比例 0.01 的增长一般就对应着实际数据 1% 的提升。对数图中的一条直线也就代表每年的百分比增长率皆为某一相同常数[1]。

以连续复利为前提，取对数操作或者总结变化的好处都远比单纯地关注百分比变化要多。比如，如果你的投资组合 (portfolio) 价值向上增长了 50%(比方说从 \$100 到 \$150)，随后又下降了相同 50%(从 \$150 到 \$75)，你并没有回到一开始的状态。如果你计算平均百分比回报 (average percentage return)(在这个例子里，0%)，这并不是一个特别有帮助的总结方式，因为其实你最终的实际结果相对于起始状态是损失了 25% 的资金。相比之下，如果你的投资在对数量度上增长了 0.5，然后再下降 0.5，你整好可以回到你初始的状态。同样地，你的投资的平均回报也整好就是你卖出所持时的对数金额减去你投资时的对数金额，除以你持有该投资的时间所获得的数值结果。

在分析经济学数据时，对数也经常是一个极有帮助的工具。譬如，这里有一张图，里面包含了从 1871 年开始的美国整体股票价格指数信息。基于这张图的量度规模，我们在前一百年上几乎看不到任何信息，然而数据在近几十年中出现了大量高幅度的波动与震荡。

---

[1]对数线性模型 $\ln P = kt + b$ 中，系数 k 表示 t 变化一个单位导致 P 变化的百分比。译者注

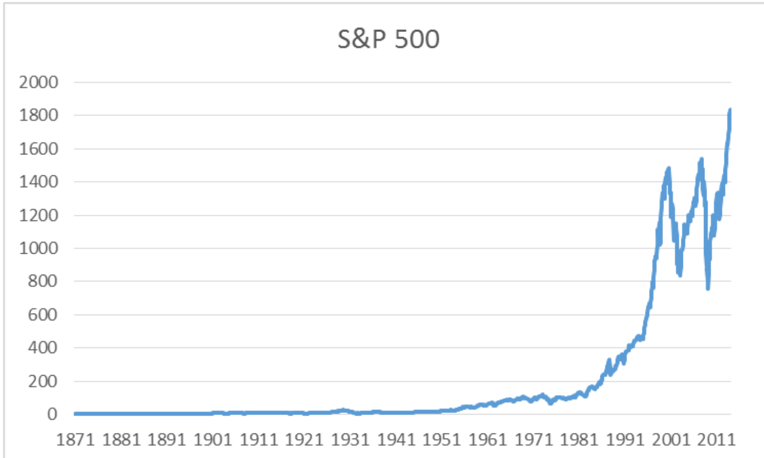

图 1: S&P 500 股票价格指数, 1871:M1 – 2014:M2. 数据来源: Robert Shiller

但另一方面，如果我们把相同的数据，用对数量度进行制图，那么纵轴上 0.01 的移动变化就等同于图中对应点 1% 的实际数值变化。通过这种方式，我们可以清晰的观察到，按百分比计算，上图中近期的股价波动实际上要比 20 世纪 30 年代所发生的震荡轻微许多。而历史上，在上个世纪 30 年代，美国正恰好处于艰难的经济大萧条时期。

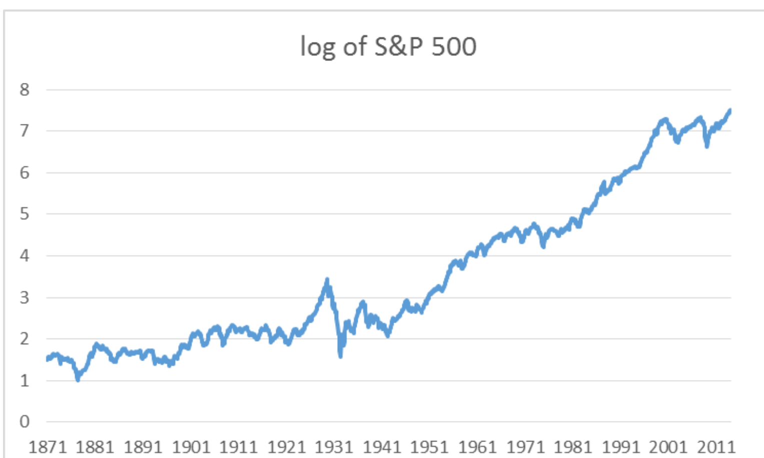

图 2: 美国股票价格指数 (自然对数) .

还有一些其他的操作变化可以帮助我们使图表更具意义且更易读。一方面，我们可能会想要剔除通货膨胀所带来的影响。下图所展示的数据我们做了一些处理，首先我们对实际股票价格而非名义股价 (nominal stock price) 取自然对数，接下来把各个结果都减去 1871 年的股价对数值，这样整个图便从 0 起始并且每个数据表示了从 1871 年开始到该年的累积对数回报 (cumulative log return)[2]。此外，所有数据都乘以了 100 倍，所以如果图中的值上升了一个单位，则对应表示股票的实际价格上升了 1%。

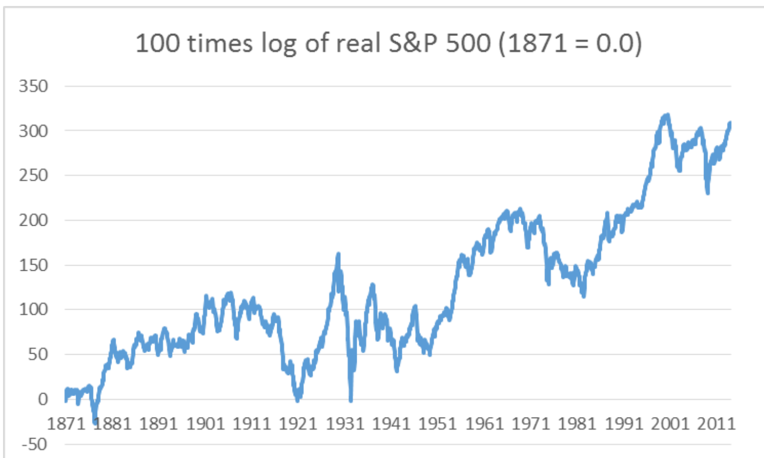

图 3: 美国股票价格指数 (自然对数数据标准化后乘以 100 倍)

自然对数在描述经济参数的关系上面也给我们提供了很大便利。举个例子，下图是一个散点图，展示了横轴上的美国实际 GDP(乘以了 100 倍并进行了标准化处理) 和纵轴上的美国石油消耗量的关系，两轴上的数据皆已采取了对数运算。这些序列之间的关系给出了一条范围约至 1970 年，斜率为 1 的直线。这条直线对应着一个值约为 1 的石油需求的收入弹性 (income elasticity) – 对于每 1% 的实际 GDP 增长，美国石油消耗量同时也上涨 1%。与此相比，从 1970 年开始，这张图的斜率变为约 0.2 – 美国实际 GDP 从那时开始年增长率 (连续复合的) 为 2.9%，但石油消耗量的增长率只为 0.6%。你能够通过计算任意两点取值的差值，再除以它们之间跨越的年份，来很容易地从图中直接获得上述的这些数据。

[2] 我们将这种操作称为对数据的标准化 (normalization)。译者注

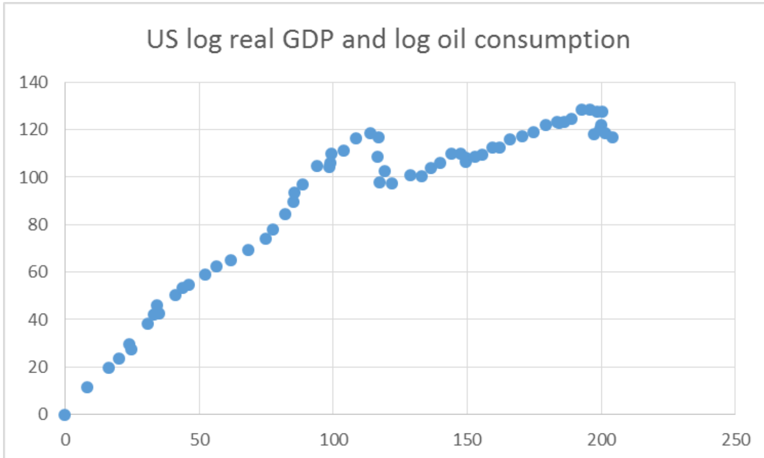

图 4: 横轴: 美国实际 GDP. 纵轴: 美国石油消耗量. 数据集覆盖范围: 1949 - 2012. 数据皆已进行取对数操作, 标准化并乘以 100 倍

所以下次如果再有人要你用对数来总结分析数据的时候，不要再抱怨了喔 ~ 这总是个能展示更多数据含义且更健壮的方法，用它来看看你的数据们到底在讲述些什么吧。