QuantNews – Fall 2007 edition

Yield curve building: interpolation versus regression

Introduction

This note discusses, in general (thus somewhat abstract) terms, issues arising while attempting to derive the interest rate structure of a particular debt market. This is a very common requirement; often a critical requirement, not only for trading in the market being analyzed, but for sales, risk management, back office, etc., as well. As a rule, such a requirement is far from easy to fulfill in a way that is satisfactory for the various purposes arising in practice.

In the vast majority of cases, the required structure takes the form of a spot or forward yield or discounting curve, rates or factors versus cash flow terms. In some cases this will be derived from pricing of closely related single and/or multiple cash flow instruments of comparable risk and liquidity. A common example would be UST on-the-run spot or forward curve built from short rates and on-the-run T-bill, Note and Bond pricing. In other cases, the curves will be derived from pricing of dissimilar but related instruments co-existing in one complex "super-market". The classic example is LIBOR curves built from short term Eurodollar borrowing-lending, Eurodollar Futures, and standard LIBOR swaps.

The TechHackers @nalyst and QuantTools products include the dfcurve family of functions to build LIBOR curves of the kind just described.

Desired properties of derived term structure curves

There are three properties that, for better or for worse, are often viewed as desirable by end product users. Perhaps unfortunately, different users often differ in the importance ascribed to these properties. We say "perhaps unfortunately", because the three properties tend to be at odds with each other; also because a solid rationale for the importance ranking may be lacking.

The property most often given priority is that the curve price back EXACTLY the instruments used to build it. (We will call this property external coherence, not standard terminology). This is quite understandable a requirement for many trading and decision support situations. We will note in passing that sometimes end users expect some rule of thumb to hold exactly when in fact it can hold only approximately with a correctly constructed curve (if at all). A frequent example is the false expectation that a derived spot rate for a given term equal the swap rate with identical term to maturity. We will not discuss this particular issue further in this technical note.

A second property often viewed as desirable (we will call it internal coherence, not standard terminology) is freedom from gross irregularities such as sudden rises or drops, cusps, unrealistic combinations of normal and inverted stretches, etc. A good rationale for wanting an internally coherent curve is the high likelihood that a "wild" stretch be an artifact of input data deficiencies. Among the most common deficiencies is the coexistence, in the inputs, of actually heterogeneous transition data directly relevant to a set of terms in close proximity (whether or not "crossovers" are actually present). The classical example of this are the transitions, in LIBOR curve building, from LIBOR borrowings / lendings to Eurodollar Futures, and from these to standard LIBOR swaps; or directly from cash to swaps.

A third property (that we might call smoothness) is the absence of angular kinks in the small. Users concerned with this want, we might say, a continuously differentiable curve over the entire range of relevant terms. We tend to believe that there is hardly ever a solid rationale for wanting this property. Other than an aesthetical one, we cannot really see a rationale that makes sense in the practice of the markets; yet, it is a goal widely judged worthy of hot pursuit. As we shall see, pursuing the aesthetical ideal is not always without its negative consequences, possibly serious ones.

We can't have everything

It should be obvious that internal coherence and external coherence fight each other. Absolute external coherence imposes a curve-building methodology that combines boot strapping with other interpolation-based calculations. Once the interpolation methods are well defined and fixed, this approach anchors the rate or discounting factor at each term in the well defined set of terms involved in the input data. Therefore, any gross irregularities in the curve can only be cured (perhaps, maybe, long shot) by improving the interpolation methods, or (most likely unavoidably) abandoning 100% external consistency. That means abandoning pure interpolation methods; and that means, in turn, adopting regression methods instead.

Theoretically, smoothness does not have to fight internal consistency when using regression methods. Pursuing smoothness does complicate regression, but does not seriously impair it. On the other hand, a quest for smoothness when using interpolation-based approaches does fight internal consistency, by imposing "advanced" interpolation methods far more prone to aggravate than to lessen bad data quirks.

The basic choices are thus: priority to external coherence, implemented via boot strapping / interpolation methods; or a compromise between internal and external coherence, implemented via regression methods. Smoothness essentially does not matter.

The TechHackers dfcurve family of functions implement a pure boot strapping / interpolation approach, with a choice of interpolation methods.