The Schwarzian Derivative & the Critical Orbit

Wes McKinney - 18.091 - 20 April 2005

1 Schwarzian Derivative

What we introduce here is a very strange looking derivative, but it turns out to give some useful information about the behavior of many functions, particularly at their critical points.

Definition: The Schwarzian derivative (SD for short) of \( f \) is

\[
S_f(x) = \frac{f'''(x)}{f'(x)} - \frac{3}{2} \left( \frac{f''(x)}{f'(x)} \right)^2.
\]

Devaney’s text has a few example of functions that we have been examining which have negative SD, and in those cases the SD at the critical points is \(-\infty\). As it will turn out, this property of having a negative SD is important.

Proposition: Let \( p(x) \) be a polynomial and suppose all of the roots of \( p'(x) \) are distinct and real. Then \( S_p < 0 \).

Proof: Since all of the roots of \( p' \) are distinct, we can factor \( p' \) as \( p'(x) = \alpha(x - \beta_1) \cdots (x - \beta_n) \), where \( \beta_1, \ldots, \beta_n \) are the \( n \) distinct roots of \( p' \), with each \( \beta_i \in \mathbb{R} \). Taking the logarithm of both sides of this equation allows us to split this product up into a summation. Differentiating kills off the constant, and dividing through by \( p' \) gives:

\[
\frac{p''(x)}{p'(x)} = \sum_{i=1}^{n} \frac{1}{x - \beta_i}.
\]

Using the quotient rule to differentiate again, we see:

\[
\frac{p'''(x)p'(x) - (p''(x))^2}{(p'(x))^2} = \frac{p'''(x)}{p'(x)} - \left( \frac{p''(x)}{p'(x)} \right)^2 = -\sum_{i=1}^{n} \frac{1}{(x - \beta_i)^2}.
\]

This gives

\[
S_p(x) = -\sum_{i=1}^{n} \frac{1}{(x - \beta_i)^2} - \frac{1}{2} \left( \sum_{i=1}^{n} \frac{1}{x - \beta_i} \right)^2 < 0.
\]

Chain Rule for Schwarzian Derivatives: Let \( f \) and \( g \) be two functions, then

\[
S(f \circ g)(x) + Sf(g(x)) \cdot (g'(x))^2 + Sg(x).
\]

Proof: Omitted, see the text.

Corollary: Suppose \( Sf < 0 \) and \( Sg < 0 \), then \( S(f \circ g) < 0 \). Along these lines, if \( Sf < 0 \), then \( Sf^n < 0 \).

Proof: This is a direct consequence of the chain rule for the SD, since \( Sf(g(x)) < 0 \) and \( Sg(x) < 0 \) \( \forall x \).

Here is one of the more interesting applications of the SD:
**Schwarzian Max-Min Principle:** Suppose \( Sf < 0 \), then \( f' \) cannot have a positive local minimum or a negative local maximum.

**Proof:** Let \( x_0 \) be a critical point of \( f' \). This means that \( f''(x_0) = 0 \). Suppose that \( f'(x_0) \neq 0 \), then

\[
Sf(x_0) = \frac{f'''(x_0)}{f'(x_0)} < 0.
\]

Now suppose \( f' \) has a positive local minimum at \( x_0 \), then by elementary calculus, \( f''(x_0) \geq 0 \), but if \( x_0 \) is a local minimum then \( f'(x_0) > 0 \), which would cause \( Sf(x_0) \geq 0 \), contradiction. Considering a negative local maximum at \( x_0 \) gives a similar contradiction.

Figure 12.1 in the text offers types of graphs which cannot occur due to this Max-Min Principle.

## 2 Critical Point and Basins of Attraction

Now we will use the SD to actually analyze the behavior of some functions. First, we need to define some things.

**Definition:** If \( x_0 \) is an attracting fixed point for a function \( f \), the **basin of attraction** of \( x_0 \) is of the set of all points whose orbits tend to \( x_0 \). The **immediate basin of attraction** is the largest interval containing \( x_0 \) that lies in the basin of attraction.

It it usually the case that the immediate basin is smaller than the full basin, but it is possible for them to be the same. If \( f(x) = x^2 \), for example, \((-1, 1)\) is the basin of attraction of 0, and it is also the immediate basin. In terms of the SD, we can now say the following:

**Theorem:** Suppose \( f \) has negative SD. If \( x_0 \) is an attracting periodic point of \( f \), then either the immediate basin of attraction of \( x_0 \) extends to \( +\infty \) or \( -\infty \), or else there is a critical point of \( f \) whose orbit is attracted to the orbit of \( x_0 \).

**Proof:** Omitted, though the text offers a proof of one of the simple cases of the theorem.

As an immediate application of this theorem to the quadratic map \( Q_c \), we note that if \(|x|\) is large enough, then the orbit tends to \( \pm \infty \). Thus, the second part of the theorem applies and we conclude that \( Q_c \) has at most one attracting cycle since \( Q_c \) has only one critical point, and \( SQ_c < 0 \).

This theorem helps explain why critical points are used to plot orbit diagrams, since the critical point will “find” the attracting cycles of the function.

**Example:** Consider \( A_\lambda(x) = \lambda \arctan x \). We have then that

\[
A'_\lambda(x) = \frac{\lambda}{1 + x^2}.
\]

If \( \lambda \neq 0 \), then \( A'_\lambda(x) \neq 0 \), so \( A_\lambda \) has no critical points. If \( |\lambda| < 1 \), then 0 is an attracting fixed point, with the immediate basin of attraction being the whole real line. If \( \lambda > 1 \), then \( A_\lambda \) has two attracting fixed points, and the basins extend to infinity. Along the same lines, when \( \lambda < -1 \), \( A_\lambda \) has an attracting 2-cycle, and the basins extent to infinity.