THE LANCZOS DERIVATIVE

LIZZY WASHBURN

ABSTRACT. The Lanczos derivative is attributed to Hungarian mathematician Cornelius Lanczos who developed it in the 1950s. It is an integral based derivative derived from the least squares model. The Lanczos derivative is set apart from other forms of differentiation such as the symmetric and traditional derivatives because it exists for functions where the other derivatives do not.

1. The Definition of the Lanczos Derivative

1.1. **Overview of the Least Squares Model.** Since the earliest times, humans have been making and recording observations of the world around them. As this process of data collection became more sophisticated, the demand for an accurate mathematical representation of this data also increased. One of the major problems mathematicians faced was the problem of error in physical observations mainly due to human and instrumental imperfections. In the beginning of the 19^{th} century, Carl Friedrich Gauss came up with a "fix" to this problem by using the argument that greater and smaller errors are equally possible in all equations. His process, which became known as the "method of least squares," found the parameters at which the sum of the squares of the errors was minimized. Today, the method remains an important statistical tool used in many fields such as business, engineering, and marketing.

Assume we have a set of discrete data: $y_1, y_2, ..., y_n$. The least squares method finds the function $f(k)$ such that

$$\sum_{k=1}^{n}(f(k) - y_k)^2$$

is a minimum. In this section we will look at how the mathematician Cornelius Lanczos used this least squares model to derive a method of differentiation.

1.2. **Finding the Derivative.** A simple way to approximate the slope between consecutive data points would be to take the change in y and divide it by the change in x. Unfortunately, any kind of error in y would greatly skew the result. Thus, it would be advantageous to use more points, the idea being that the slope changes little over the span of consecutive points. Although one can look at any number of points at a time, we will start by looking at five points y_{-2}, y_{-1}, y_0, y_1, and y_2, and assume they lie closely to a parabola. Therefore, let's use the formula

$$f(k) = a + bk + ck^2,$$

where a, b, c are coefficients which can be adjusted for our data. Using the least squares method, we set up the function

$$\sum_{k=-2}^{2} (f(k) - y_k)^2 .$$

By minimizing the sum of these squares with respect to a, b, c, we will find the best fit parabola. Thus, the function we want to minimize looks like

$$\sum_{k=-2}^{2} \left[(a + bk + ck^2) - y_k\right]^2 . \tag{1}$$

Since we are trying to find the value of the derivative of at $x = 0$, we can differentiate $f(k) = a + bk + ck^2$ and get $f'(0) = b$. Hence it is the constant b we are interested in. The condition of minimum with respect to b gives us

$$\begin{aligned} 0 &= -4(a - 2b + 4c - y_{-2}) - 2(a - b + c - y_{-1}) \\ &+ 2(a + b + c - y_1) + 4(a + 2b + 4c - y_2). \end{aligned}$$

All the constants except for b cancel out resulting in

$$b = \frac{-2y_{-2} - y_{-1} + y_1 + 2y_2}{10}.$$

So far we have worked under the assumption that we have five data points. Let's assume instead that the number of data points is $2n+1$, which would therefore give us a more general equation for approximating the derivative. The sum in (1) is the function which we then differentiated in terms of b and set equal to 0. By doing this, we found our approximate derivative b. By doing this same process, only with $2n+1$ data points instead of five, we get

$$\sum_{k=-n}^{n} 2k\left[(a + bk + ck^2) - y_k\right] = 0.$$

Since again, the coefficients a and c drop out, our equation becomes

$$\begin{aligned} 0 &= \sum_{k=-n}^{n} 2k(bk - y_k) \\ \sum_{k=-n}^{n} 2k^2 b &= \sum_{k=-n}^{n} 2ky_k \\ b &= \frac{\sum_{k=-n}^{n} ky_k}{2\sum_{k=1}^{n} k^2}. \end{aligned}$$

Here we have assumed that our interval between consecutive observations is one. If we make that interval instead h, we must divide our derivative by the length h and the formula of differentiation for our empirical function becomes

$$b = \frac{\sum_{k=-n}^{n} kf(x+kh)}{2\sum_{k=1}^{n} k^2 h}. \tag{2}$$

Thus far we have looked only at functions with discrete variables. Let's assume that we have a dense group of observations, each observation very close to the next. Since we are looking at a very dense group of observations, our sum becomes an integral. We can then rewrite b as

$$b = \frac{\int_{-\epsilon}^{\epsilon} tf(x+t)dt}{\frac{2\epsilon^3}{3}} = \frac{3}{2\epsilon^3}\int_{-\epsilon}^{\epsilon} tf(x+t)dt,$$

where ϵ is sufficiently small. From here, Lanczos derived the formula for differentiation

$$f_L'(x) = \lim_{\epsilon\to 0}\frac{3}{2\epsilon^3}\int_{-\epsilon}^{\epsilon} tf(x+t)dt. \tag{3}$$

This is the definition of the Lanczos derivative.[2]

1.3. **Illustration of the Derivative.**

Let's do an example and find the Lanczos derivative of $f(x) = x^3$. Using the formula, we get

$$\begin{aligned}
f_L'(x) &= \lim_{\epsilon\to 0}\frac{3}{2\epsilon^3}\int_{-\epsilon}^{\epsilon} t(x+t)^3 dt \\
&= \lim_{\epsilon\to 0}\frac{3}{2\epsilon^3}\Big(\frac{1}{2}t^2x^3 + x^2t^3 + \frac{3}{4}xt^4 + \frac{1}{5}t^5\Big)\Big|_{-\epsilon}^{\epsilon} \\
&= \lim_{\epsilon\to 0}\Big(3x^2 + \frac{3}{5}\epsilon^2\Big) \\
&= 3x^2
\end{aligned}$$

which is, as we should expect, the traditional derivative of x^3.

2. The Accuracy of the Lanczos Derivative

Looking at the error of the Lanczos derivative sheds light on both the basic uses and the limitations of the derivative.

2.1. **The Lanczos Derivative and the Taylor Series.**

Let's assume we can use the Taylor series to represent the function in the Lanczos formula $f(x+t)$. Recall that the Taylor series representation of $f(x)$ centered at $x = a$ is

$$f(x) = f(a) + \frac{f'(a)(x-a)}{1!} + \frac{f''(a)(x-a)^2}{2!} + \frac{f^{(3)}(a)(x-a)^3}{3!} + \cdots .$$

Let's choose x as the center of $f(x+t)$ where x is fixed and t varies. Therefore, the Taylor series representation looks like

$$f(x+t) = f(x) + \frac{t}{1!}f'(x) + \frac{t^2}{2!}f''(x) + \frac{t^3}{3!}f^{(3)}(x) + \cdots .$$

We can then find the Lanczos derivative using this sum

$$\begin{aligned} f_L'(x) &= \lim_{\epsilon \to 0} \frac{3}{2\epsilon^3} \int_{-\epsilon}^{\epsilon} t f(x+t) dt \\ &= \lim_{\epsilon \to 0} \frac{3}{2\epsilon^3} \left(\frac{1}{2} \cdot t^2 f(x) + \frac{1}{3} \cdot \frac{t^3}{1!} f'(x) + \frac{1}{4} \cdot \frac{t^4}{2!} f''(x) + \cdots \right) \Big|_{-\epsilon}^{\epsilon} . \end{aligned}$$

Looking at the Lanczos derivative as an infinite sum can give us some interesting insight into how the Lanczos derivative behaves. If we look at this integral, we can see that all the even powers of t are going to cancel each other out when we substitute $\pm\epsilon$ in for t while our odd powers of t double. Therefore, our derivative becomes

$$f_L'(x) = \lim_{\epsilon \to 0} \left(f'(x) + \frac{1}{10}\epsilon^2 f^{(3)}(x) + \cdots \right).$$

Thus, the Lanczos derivative does in fact equal $f'(x)$ with an error of the order ϵ^2. In Lanczos' *Applied Analysis*, he talks about how this error, which he calls "noise," typically destroys the analytical nature of $f(x)$.[2] Yet Lanczos argues that this error is greatly diminished by the randomness of the errors, which through summation, balances itself out.

2.2. **A closer look at the Error of the Lanczos Derivative.** In *Lanczos' Generalized Derivative*,[3] C. W. Groetsch analyzes more closely the error of the Lanczos derivative. Small errors in the function values are magnified by a factor of $\frac{3}{2\epsilon}$. This, Groetsch states, "is not a deficiency in the method, but rather a manifestation of the inherent instability of the differentiation process itself."[3] To look at this more closely, let us allow $f^h(t)$ to be some bounded integrable perturbation of f where $|f(t) - f^h(t)| \leq h$ and h is a known error bound. If we let

$f'_L(t)$ be the Lanczos derivative, then

$$|\left(f^h\right)'_L(x) - f'(x)| \leq |f'_L(x) - f'(x)| + |\left(f^h\right)'_L(x) - f'_L(x)|.$$

Earlier we found that the error of the Lanczos derivative was $\frac{\epsilon^2}{10} f^{(3)}(x)$ by using the Taylor series representation. Let us assume $f^{(3)}$ exists and is bounded by M. Then we can say

$$|f'_L(x) - f'(x)| \leq \frac{M}{10}\epsilon^2.$$

We can also see that

$$\begin{aligned} \left|\left(f^h\right)'_L(x) - f'_L(x)\right| &= \left|\frac{3}{2\epsilon^3}\int_{-\epsilon}^{\epsilon} tf^h(x+t) - tf(x+t)dt\right| \\ &\leq \frac{3}{2\epsilon^3}\int_{-\epsilon}^{\epsilon} |t|\, hdt = \frac{3h}{2\epsilon}. \end{aligned}$$

We can then find our total error by adding these two together, and we get

$$\left|\left(f^h\right)'_L(x) - f'(x)\right| \leq \frac{M}{10}\epsilon^2 + \frac{3h}{2\epsilon}.$$

Groetsch calls the first error the truncation error because it goes to zero as $\epsilon \to 0$. The second error is called the stability error; it blows up as $\epsilon \to 0$. Therefore, it's necessary to find a balance between the two errors to stay between 'the two numerical hazards.'

3. The Traditional and Symmetric Derivatives

Before we go further into our analysis of the Lanczos derivative, it makes sense to compare it to other forms of differentiation with which we are probably more familiar. In this section we will look at the traditional and symmetric derivatives and see how their properties compare to the Lanczos derivative.

3.1. The Traditional Derivative.

The traditional derivative finds the slope by looking at the ordered pairs $(x, f(x))$ and $(x+h, f(x+h))$ and finding the rise over run between these points as $h \to 0$:

$$f'(x) = \lim_{h\to 0}\frac{f(x+h) - f(x)}{h}. \tag{4}$$

A function is differentiable at a point x if the limit exists. Otherwise, the function is not differentiable.

3.2. The Symmetric Derivative.

The symmetric derivative is similar to the traditional derivative. The difference between the two is that the symmetric derivative

$$f_s'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x-h)}{2h} \tag{5}$$

uses the points $(x-h, f(x-h))$ and $(x+h, f(x+h))$ which lie on either side of $(x, f(x))$ while the traditional derivative uses the point $(x, f(x))$ as one of its two points of interest.

3.3. Comparison of the Traditional and Symmetric Derivatives.

Theorem 3.1. *If $f'(x)$ exists, then $f_s'(x)$ exists.*

Proof. Let $f(x)$ be differentiable at $x = a$. Therefore,

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

exists. The symmetric derivative can be rewritten as

$$\begin{aligned} f_s'(a) &= \lim_{h \to 0} \frac{f(a+h) - f(a-h)}{2h} \\ &= \lim_{h \to 0} \left(\frac{f(a+h) - f(a)}{2h} + \frac{f(a) - f(a-h)}{2h} \right). \end{aligned}$$

Since $f'(a)$ exists, then

$$\frac{1}{2} \cdot \frac{f(a+h) - f(a)}{h} + \frac{1}{2} \cdot \frac{f(a + (-h)) - f(a)}{(-h)}$$

has a limit of $f'(a)$ as $h \to 0$. Therefore, $f_s'(a) = f'(a)$ also exists. □

It's important to note that if $f_s'(x)$ exists, $f'(x)$ need not exist.[4] One example of this is the derivative of $f(x) = |x|$ at $x = 0$. With the traditional derivative, we come to a problem:

$$f'(0) = \lim_{h \to 0} \frac{|0 + h| - |0|}{h} = \lim_{h \to 0} \frac{|h|}{h}$$

The limit does not exist, and therefore the function is not differentiable at $x = 0$. Under $f_s'(x)$, we get a different outcome:

$$f_s'(0) = \lim_{h\to 0} \frac{|0+h| - |0-h|}{2h} = \lim_{h\to 0} \frac{0}{2h} = 0.$$

Therefore, the derivative under this method exists and equals 0. The symmetric derivative gives us another method to compute the slope when the traditional one does not work.

3.4. **Comparison of the Symmetric and Lanczos Derivatives.** When we were constructing the Lanczos derivative, we computed formula (2)

$$\frac{\displaystyle\sum_{n=-k}^{k} nf(x+nh)}{\displaystyle 2\sum_{n=1}^{k} n^2 h}.$$

We were able to obtain the Lanczos derivative from (2) by assuming that our interval h became infinitely small as our number of points n became infinitely dense. If we look at (2) when $n = 1$, our function looks surprisingly familiar. We obtain

$$\frac{-f(x-h) + f(x+h)}{2h}.$$

Letting $h \to 0$, we get the exact definition of the symmetric derivative. If we let $n = 2$, our new equation becomes

$$\frac{-2f(x-2h) - f(x-h) + f(x+h) + 2f(x+2h)}{10h}.$$

With a little mathematical manipulation, we can rewrite this as

$$\frac{4}{5} \cdot \frac{f(x+2h) - f(x-2h)}{4h} + \frac{1}{5} \cdot \frac{f(x+h) - f(x-h)}{2h} = \frac{4}{5} f_s' + \frac{1}{5} f_s' = f_s'.$$

This trend continues. If we were to continue to compute the slope for varying values of n, we would see that

$$\frac{\displaystyle\sum_{n=-k}^{k} nf(x+nh)}{\displaystyle 2\sum_{n=1}^{k} n^2 h} = \frac{1^2}{\displaystyle\sum_{n=1}^{k} n^2} f'_s + \frac{2^2}{\displaystyle\sum_{n=1}^{k} n^2} f'_s + \cdots + \frac{k^2}{\displaystyle\sum_{n=1}^{k} n^2} f'_s \;=\; f'_s.$$

This sheds a new light on the Lanczos derivative. The symmetric derivative looks at this tiny interval on each side of x and averages the slope of each side. Lanczos uses the same idea, but uses an infinite number of these minute intervals to measure the gradient. His end outcome is that he uses integration to find its opposite process: differentiation. This seems strange initially, but when we gain an understanding of what we are actually computing, it makes complete sense. The symmetric derivative works by finding the slope over an interval and then letting the interval shrink to 0 whereas the Lanczos derivative works by finding the slope over many intervals, and then shrinking each of these intervals.

Theorem 3.2. *If $f'_S(x)$ exists, then $f'_L(x)$ exists.*

Proof. Recall that the definition of the Lanczos derivative is

$$f'_L(x) = \lim_{\epsilon\to 0} \frac{3}{2} \cdot \frac{\int_{-\epsilon}^{\epsilon} tf(x+t)dt}{\epsilon^3}.$$

This is of the $\frac{0}{0}$ indeterminate form as $\epsilon \to 0$, so if we assume that f is continuous, we can apply L'Hospital's Rule and the Fundamental Theorem of Calculus to the limit to get

$$f'_L(x) = \lim_{\epsilon\to 0} \frac{3}{2} \cdot \frac{\epsilon f(x+\epsilon) - \epsilon f(x-\epsilon)}{3\epsilon^2} \;=\; \lim_{\epsilon\to 0} \frac{f(x+\epsilon) - f(x-\epsilon)}{2\epsilon} = f'_s(x).$$

□

Therefore, if f is continuous, the limit of the Lanczos derivative equals the symmetric derivative and therefore the Lanczos derivative exists everywhere the symmetric derivative does.

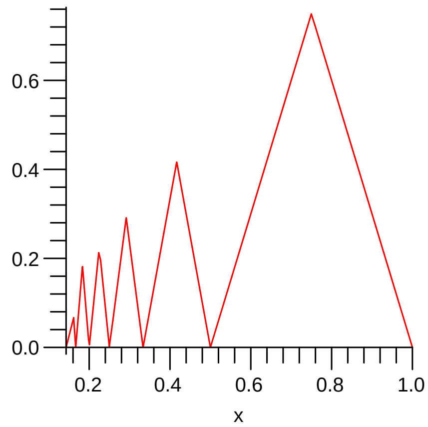

FIGURE 1. The the graph of $y = f(x)$ is a sequence of isosceles triangles which grow as $x \to \infty$

3.5. **The Relationship of the Derivatives.**

Conjecture 3.1. *Let TD, SD, and LD denote the class of continuous functions with a traditional derivative, symmetric derivative, and Lanczos derivative respectively. Then*

$$TD \subseteq SD \subseteq LD$$

We have already shown that if the traditional derivative exists, then so too do the symmetric and therefore the Lanczos derivatives. We also saw that if the symmetric derivative exists, so does the Lanczos derivative. This is not a two way relationship, though. We have looked at an example where the symmetric derivative exists, but the traditional derivative does not. Is there an example where the Lanczos derivative exists, but the symmetric derivative (and hence also the traditional derivative) does not exist? This example is more difficult to construct and is considered in the next section.

4. A FUNCTION WITH SOLELY A LANCZOS DERIVATIVE

The function $f(x)$ shown in Figure 1 equals 0 when $x < 0$ and is comprised of isosceles triangles when $x \geq 0$. Two corners of each triangle lie on the x-axis and we'll label these points: $r = \frac{1}{n}$ and $l = \frac{1}{n+1}$. The upper point of the triangle lies on the line $y = x$ at the point (c, c) where c is the average of r and l. As $x \to 0$, the triangles get infinitely small and the number of triangles approaches infinity. The equation of our function is

$$f(x) = -\frac{c}{c-l}\,|x-c| + c \quad \text{for} \quad l \leq x < r.$$

For $x > 0$ and x not lying on any corner of a triangle, we can fairly easily find the traditional derivative $f'(x)$. We can simply treat the function like a piece-wise function and find the derivative of the linear equation over the interval containing the point of interest. The symmetric derivative would also exist for these points x, as well as the points at the triangle corners, the derivative in this case being the average slope of either side of the point. Therefore, the symmetric derivative exists for each point $x > 0$. At $x = 0$, though, the symmetric derivative no longer exists. Imagine a line going through the upper corner of each of the triangles. The equation of this line is $y = x$ and therefore has a slope of 1. A line going through the lower corners of each of the triangles has a slope of 0. These contrasting values mean that the limit of the symmetric derivative does not exist.

We do not run into this same problem with the Lanczos derivative. We can set up the Lanczos derivative with

$$f_L'(0) = \lim_{k\to\infty} \frac{3k^3}{2} \sum_{n=k}^{\infty} \int_l^r t\left[-\frac{c}{c-l}\,|t-c| + c\right]\,dt$$

where $r = \frac{1}{n}$, $l = \frac{1}{n+1}$, and $c = \frac{2n+1}{2n(n+1)}$. Our equation within the integral is $f(x+t)$ when $x = 0$. Above, we have integrated over the interval of one triangle, and then found the sum for the infinite number of triangles, starting at $x = \frac{1}{k}$ and letting $k \to \infty$. Since in this case, $\epsilon = \frac{1}{n}$, we have to make the proper adjustments to our equation.

We can evaluate the integral by changing our intervals so the absolute value can be removed:

$$\begin{aligned}\int_l^r t\left[-\frac{c}{c-u}\left|t-c\right|+c\right]\,dt \;&=\; \int_c^r t\left[-\frac{c}{c-u}(t-c)+c\right]\,dt\\ &+\;\int_l^c t\left[\frac{c}{c-u}(t-c)+c\right]\,dt.\end{aligned}$$

Then, we can rewrite the Lanczos derivative in terms of n to get

$$\begin{aligned}\lim_{k\to\infty}\frac{3k^3}{2}\sum_{n=k}^{\infty}&\left[\int_{\frac{2n+1}{2n(n+1)}}^{\frac{1}{n}} t\left[-(2n+1)\left(t-\frac{2n+1}{2n(n+1)}\right)+\frac{2n+1}{2n(n+1)}\right]\,dt\right.\\ &\left.+\int_{\frac{1}{n+1}}^{\frac{2n+1}{2n(n+1)}} t\left[(2n+1)\left(t-\frac{2n+1}{2n(n+1)}\right)+\frac{2n+1}{2n(n+1)}\right]\,dt\right].\end{aligned}$$

By evaluating the integrals, we obtain

$$f_L'(0)=\lim_{k\to\infty}\frac{3k^3}{2}\sum_{n=k}^{\infty}\frac{(2n+1)^2}{8n^3(n+1)^3}.$$

In order to evaluate $\sum_{n=k}^{\infty}\frac{(2n+1)^2}{8n^3(n+1)^3}$, let us compare the sum and the integral of $\frac{(2x+1)^2}{8x^3(x+1)^3}$. Since $y=\frac{(2x+1)^2}{8x^3(x+1)^3}$ is a decreasing function when $x>0$,

$$\frac{3k^3}{2}\int_k^{\infty}\frac{(2x+1)^2}{8x^3(x+1)^3}dx<\frac{3k^3}{2}\sum_{n=k}^{\infty}\frac{(2n+1)^2}{8n^3(n+1)^3}.$$

Also, the fact that y is a decreasing function means that

$$\frac{3k^3}{2}\sum_{n=k+1}^{\infty}\frac{(2n+1)^2}{8n^3(n+1)^3}<\frac{3k^3}{2}\int_k^{\infty}\frac{(2x+1)^2}{8x^3(x+1)^3}dx$$

or

$$\frac{3k^3}{2}\sum_{n=k}^{\infty}\frac{(2n+1)^2}{8n^3(n+1)^3}<\frac{3k^3}{2}\left[\frac{(2k+1)^2}{8k^3(k+1)^3}+\int_k^{\infty}\frac{(2x+1)^2}{8x^3(x+1)^3}dx\right]$$

Evaluating the integrals, we obtain

$$\frac{1}{4}<\frac{3k^3}{2}\sum_{n=k}^{\infty}\frac{(2n+1)^2}{8n^3(n+1)^3}<\frac{3(2k+1)^2}{16(k+1)^3}+\frac{1}{4}.$$

Taking the limit as $k \to \infty$ of this inequality, we get

$$\frac{1}{4} \le \lim_{k\to\infty} \frac{3k^3}{2} \sum_{n=k}^{\infty} \frac{(2n+1)^2}{8n^3(n+1)^3} \le \frac{1}{4}.$$

Now we can use the Squeeze Theorem to find the limit of the sum.

Theorem 4.1 (Squeeze Theorem)**.** *If $f(x) \le g(x) \le h(x)$ when x is near a (except possibly at a) and*

$$\lim_{x\to a} f(x) = \lim_{x\to a} h(x) = L$$

then

$$\lim_{x\to a} g(x) = L$$

Therefore, from the Squeeze Theorem

$$\lim_{k\to\infty} \frac{3k^3}{2} \sum_{n=k}^{\infty} \frac{(2n+1)^2}{8n^3(n+1)^3} = \frac{1}{4} = \lim_{k\to\infty} \frac{3k^3}{2} \int_k^{\infty} \frac{(2x+1)^2}{8x^3(x+1)^3}\, dx.$$

We can now find the limit of the Lanczos derivative through the integral test and L'Hospital's rule:

$$\begin{aligned} f_L'(0) &= \lim_{x\to\infty} \frac{3}{2} \cdot \frac{\int_x^{\infty} \frac{(2t+1)^2}{8t^3(t+1)^3}\, dt}{\frac{1}{x^3}} = \lim_{x\to\infty} \frac{3}{2} \cdot \frac{\frac{-(2x+1)^2}{8x^3(x+1)^3}}{\frac{-3}{x^4}} \\ &= \lim_{x\to\infty} \frac{3}{2} \cdot \frac{(2x+1)^2 x}{24(x+1)^3} = \frac{1}{4}. \end{aligned}$$

Therefore, the Lanczos method can find a derivative of our function of $\frac{1}{4}$, while the symmetric and traditional derivatives do not exist. The value $\frac{1}{4}$ comes from the process of measuring the average derivative from, in this case, 0 to ϵ, over a dense number of intervals and then letting $\epsilon \to 0$.

5. Conclusion

Although Lanczos' formula is a more tedious method for finding the derivative than the other more common definitions of differentiation, the Lanczos derivative does carry some advantages over the other formulas. As we saw with the function of triangles, there are circumstances where other derivatives do not exist, but the Lanczos derivative does. This allows us to compute a 'pseudo-derivative' at places that are not differentiable under the traditional derivative.

Overall, though, the significance is not in the ability of the derivative to compute a pseudo-derivative, but rather in the properties the derivative shares with other mathematical subjects. We looked at how the Lanczos derivative compares to other derivatives and the least squares method. Other topics that can be looked at in relation to the Lanczos derivative are higher order derivatives and Legendre polynomials. For example, we can use the least squares method to look at the Lanczos derivative function $f(x+t)$. The equation

$$\int_{-\epsilon}^{\epsilon} (f(x+t) - g(x))^2 \, dt$$

where $g(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x^1 + a_0$ can be minimized, similar to the least squares process. The result we obtain by differentiating in terms of the highest order coefficient ends up being very similar to the n^{th} order Legendre polynomial. There is much more that can be researched in this area that might explain why and how this phenomenon occurs.

References

[1] Bjorck, Ake. Numerical Methods for Least Squares Problems. Philadelphia: Society for Industrial and Applied Mathematics, 1996.

[2] Lanczos, Cornelius. Applied Analysis. Englewood Cliffs, N.J.: Prentice Hall, 1961.

[3] Groetsch, C. W. Lanczos' Generalized Derivative. *The American Mathematical Monthly*; Apr 1998.

[4] Poodiac, Robert P. The Symmetric Derivative and the Unit Circle 16 March 2004, http://www2.norwich.edu/rpodiac/cmjarticle2.pdf.