Similarly, using $\{y_i\}$ for weights, the weighted mean population density is

$$\bar{d}_{yw} = \frac{\sum \frac{x_i}{y_i} y_i}{\sum y_i} = \frac{\sum x_i}{\sum y_i},$$

and the corresponding weighted variance is

$$\sigma_{yw}^2 = \frac{\sum \left(\frac{x_i}{y_i} - \bar{d}_{yw}\right)^2 y_i}{\sum y_i} = \frac{\sum \frac{x_i{}^2}{y_i}}{\sum y_i} - \bar{d}_{yw}^2,$$

a quantity that is clearly nonnegative.

Dividing the last identity by $\bar{d}_{yw}$, we have this unifying theorem:

Theorem 3. *If $\{x_i\}$ and $\{y_i\}$ are positive sequences of length n, then*

$$\bar{d}_{xw} - \bar{d}_{yw} = \sigma_{yw}^2 / \bar{d}_{yw}.$$

Thus, the difference between $\bar{d}_{xw}$ and $\bar{d}_{yw}$ is directly proportional to the variability of the county densities, and inversely proportional to $\bar{d}_{yw}$. Also note that Theorem 2 is a corollary of Theorem 3. Moreover, letting $y_i = 1$ for each i produces Hemenway's identity, and so Theorem 1 also is a corollary of Theorem 3.

Conclusion. Choosing the most useful descriptors of a data set requires careful consideration, and if it also leads one to learn classical results by two pioneers in the theory of inequalities, so much the better.

Acknowledgment. The author thanks Professor Vedula N. Murty for his helpful suggestions and advice on this topic.

References

1. D. Hemenway, Why your classes are larger than "average," *Math. Mag.* **55** (1982) 162–164.
2. A. Lann and R. Falk, A closer look at a relatively neglected mean, *Teaching Stat.* **27** (2005) 76–80.
3. D. S. Mitrinović, *Analytic Inequalities*, Springer-Verlag, New York, 1970.
4. A. Schwenk, Distortion of average class size: The Lake Wobegon effect, this JOURNAL **37** (2006) 293–296.
5. U.S. Census Bureau, *Census 2000, Summary File 1 (SF 1)*, generated by the author using American Factfinder; available at `http://factfinder.census.gov/` (24 August 2006).

Differentiating the Arctangent Directly

Eric Key (ericskey@uwm.edu), University of Wisconsin-Milwaukee, Milwaukee, WI 53201-0413
doi:10.4169/193113409X458750

We will show by a direct argument that the inverse tangent function is differentiable at all values in its domain.

Suppose, then, that $0 \le x < y$ and let

$$\theta = \arctan(y) - \arctan(x).$$

Recalling that

$$\begin{aligned}
\sin(\arctan(u)) &= \frac{u}{\sqrt{1+u^2}}, \\
\cos(\arctan(u)) &= \frac{1}{\sqrt{1+u^2}}, \\
\sin(a-b) &= \sin(a)\cos(b) - \sin(b)\cos(a), \\
\tan(a-b) &= \frac{\tan(a) - \tan(b)}{1 + \tan(a)\tan(b)},
\end{aligned}$$

we see that

$$\frac{y-x}{\sqrt{1+x^2}\sqrt{1+y^2}} \le \arctan(y) - \arctan(x) \le \frac{y-x}{1+xy},$$

since if $0 \le \theta < \pi/2$, $\sin(\theta) \le \theta \le \tan(\theta)$. Therefore, since $0 \le x < y$, we have

$$\begin{aligned}
\frac{1}{1+y^2} &< \frac{1}{\sqrt{1+x^2}\sqrt{1+y^2}} \\
&\le \frac{\arctan(y) - \arctan(x)}{y-x} \\
&\le \frac{1}{1+xy} \\
&\le \frac{1}{1+x^2}.
\end{aligned}$$

It then follows from the Pinching (or Squeeze) Principle that the inverse tangent function is differentiable on $(0, \infty)$ and has a righthand derivative at 0. Since the inverse tangent function is an odd function, this is sufficient to prove that it is differentiable on the whole real line and that

$$\frac{d}{dx}\arctan(x) = \frac{1}{1+x^2}. \tag{1}$$

In light of the identities

$$\begin{aligned}
\operatorname{arccot}(x) &= \frac{\pi}{2} - \arctan(x), \\
\arcsin(x) &= \arctan\left(\frac{x}{\sqrt{1-x^2}}\right), \\
\arccos(x) &= \frac{\pi}{2} - \arcsin(x),
\end{aligned}$$

$$\operatorname{arcsec}(x) = \arccos\left(\frac{1}{x}\right),$$

$$\operatorname{arccsc}(x) = \frac{\pi}{2} - \operatorname{arcsec}(x),$$

(1) implies that each inverse trigonometric function is differentiable on the interior of its domain. Consequently, each derivative can then be computed by implicit differentiation or the chain rule.

Finding Matrices that Satisfy Functional Equations

Scott Duke Kominers (skominers@gmail.com), Harvard University
doi:10.4169/193113409X458769

Linear algebra is typically taught shortly after calculus. Thus, calculus-informed linear algebra problems offer an exceptional opportunity to illustrate interactions between different branches of undergraduate mathematics.

For example, consider the following problem, which (as we shall see) has its roots in calculus:

Problem. Find a 4×4, nonsingular, nonconstant matrix function $N(x)$ that satisfies the functional equation

$$N(2x) - (N(x))^8 = 0.$$

At first glance, this problem appears to be quite difficult. Beyond the likely difficulty of finding such a matrix $N(x)$, it is not even immediately clear how one would *prove* that a matrix $N(x)$ is actually a solution without a great deal of matrix algebra. However, this problem is not hard as it seems. In fact, it is one of a large class of problems that can be solved via a surprising method based upon single-variable calculus.

In this JOURNAL, Khan [**2**] used nilpotent matrices and Taylor series to find matrix functions satisfying the exponential functional equation, $f(x+y) = f(x) \cdot f(y)$. His method is an example of a much more general theory of matrix power series due to Weyr [**4**], which can be used to find matrix functions satisfying a variety of functional equations. (Rinehart [**3**] gives an excellent survey of Weyr's approach. Higham [**1**, ch. 4] gives a more comprehensive account, as well as further applications.)

We say that a set of real-valued functions $\{f_i(x)\}_{i=1}^n \subset \mathcal{C}^\infty(\mathbb{R})$ *satisfies an analytic functional equation* E if there is an analytic function E such that

$$E(f_1, \ldots, f_n)(x) = 0$$

identically for all $x \in \mathbb{R}$. For example, the trigonometric functions $f_1(x) = \sin(x)$ and $f_2(x) = \cos(x)$ satisfy the analytic functional equation

$$E(f_1, f_2) = f_1^2 + f_2^2 - 1 \equiv 0.$$

Now, for any set of real-valued functions $\{f_i(x)\}_{i=1}^n \subset \mathcal{C}^\infty(\mathbb{R})$ satisfying the analytic functional equation E, we will find a set of associated matrix functions $\{A_i(x)\}_{i=1}^n$ satisfying the same functional equation E.