

Introduction to



for statistics and probability

**Cornell Statistics Graduate Society**

Sara Venkatraman

# Outline

1. What is Mathematica and why should you use it?
2. How to obtain Mathematica
3. Live demo
4. Example of Mathematica in research

# What is Mathematica?

- Mathematica is a **symbolic** computing system: can manipulate mathematical expressions, unlike R or Python
- Uses the Wolfram programming language
- Example: let  $f(x, \theta) = (\theta x^3 + 1)^{-1}$ . What is  $\partial f / \partial x$ ?

WOLFRAM MATHEMATICA | STUDENT EDITION

```
In[1]:= f[x_,  $\theta$ _] := ( $\theta$  x^3 + 1) ^ (-1)
```

```
In[2]:= D[f[x,  $\theta$ ], x]
```

```
Out[2]= 
$$-\frac{3 x^2 \theta}{(1 + x^3 \theta)^2}$$

```

# What is Mathematica?

For statisticians, I think Mathematica is very useful for:

- Verifying tedious calculus: derivatives, integrals, etc.
- Analytically solving complicated equations
- Working with random variables (not realizations of them)

I think Mathematica is less useful for:

- Working with and visualizing real data
- Larger and potentially open-source software

# How to get Mathematica

- \$35 for one-year license (Aug. 1 - July 31) from Cornell IT; also includes Mathematica Online and WolframAlpha Pro:  
<https://it.cornell.edu/software-licensing/mathematica-licensing>
- 15-day free trial:  
<https://www.wolfram.com/mathematica/trial/>

## Live Mathematica demo

# Mathematica in my research

A Bayesian approach to regression:

- Consider the model  $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ , with  $\varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ .
- Put a  $N(\beta_0, \sigma^2 \mathbf{V}_0)$  prior on  $\beta$ , where  $\mathbf{V}_0$  is some p.s.d. matrix.
- What should  $\mathbf{V}_0$  be? Let's say  $\mathbf{V}_0 = g(\mathbf{X}^T \mathbf{X})^{-1}$ , for some  $g > 0$
- Then the posterior mean of  $\beta$  is:

$$\mathbb{E}(\beta | \mathbf{Y}, \sigma^2) = \frac{1}{1+g} \beta_0 + \frac{g}{1+g} \hat{\beta}$$

where  $\hat{\beta}$  is the usual least-squares estimate!

- **What I want to know: How do we choose  $g$ ?**

## Mathematica in my research (continued)

We could try solving for  $g$  that minimizes sum of squared residuals,

$$\|\mathbf{Y} - \hat{\mathbf{Y}}\|^2, \text{ where } \hat{\mathbf{Y}} = \mathbf{X}(\frac{1}{1+g}\beta_0 + \frac{g}{1+g}\hat{\beta}).$$

After some algebra, I can write this as:

$$\text{SSR}(g) = \|\mathbf{Y} - \hat{\mathbf{Y}}\|^2 = a - \frac{2b}{1+g} - \frac{2gc}{1+g} + \frac{d}{(1+g)^2} + \frac{2gb}{(1+g)^2} + \frac{g^2c}{(1+g)^2}$$

where  $a = \|\mathbf{Y}\|^2$ ,  $b = \mathbf{Y}^T \mathbf{Y}_0$ ,  $c = \|\hat{\beta}\|^2$ ,  $d = \|\mathbf{Y}_0\|^2$ ,  $\mathbf{Y}_0 = \mathbf{X}\beta_0$ .

```
In[1]:= SSR[g_] := a - 2 b / (1 + g) - 2 g c / (1 + g) + d / (1 + g) ^ 2 +  
          2 g b / (1 + g) ^ 2 + g ^ 2 c / (1 + g) ^ 2  
  
In[2]:= Solve[D[SSR[g], g] == 0, g]  
  
Out[2]= {}
```

Turns out there are no minimizing values of  $g$ ...



## Mathematica in my research (continued)

Instead let's minimize *Stein's unbiased risk estimate* (SURE), given by:

$$\|\mathbf{Y} - \mathbf{X}\hat{\beta}\|^2 + \frac{2gp\hat{\sigma}^2}{1+g} - n\hat{\sigma}^2.$$

After some algebra, I can write this as:

$$\text{SURE}(g) = a - \frac{2b}{1+g} - \frac{2gc}{1+g} + \frac{d}{(1+g)^2} + \frac{2gb}{(1+g)^2} + \frac{g^2c}{(1+g)^2} + \frac{2gp\hat{\sigma}^2}{1+g} - n\hat{\sigma}^2$$

```
In[1]:= SURE[g_] := a - 2 b / (1 + g) - 2 g c / (1 + g) + d / (1 + g) ^ 2 +  
        2 g b / (1 + g) ^ 2 + g ^ 2 c / (1 + g) ^ 2 + 2 g p σ ^ 2 / (1 + g) -  
        n σ ^ 2
```

```
In[2]:= Solve[D[SURE[g], g] == 0, g]
```

```
Out[2]= {{g ->  $\frac{-2 b + c + d - p \sigma^2}{p \sigma^2}$ }}
```

Now we have a solution; let's check that it's indeed a minimizer.

## Mathematica in my research (continued)

Take second derivative of SURE with respect to  $g$       Evaluate second derivative at the optimum      Simplify resulting expression

In[5]:=  $D[\text{SURE}[g], \{g, 2\}] /. g \rightarrow \frac{-2b + c + d - p\sigma^2}{p\sigma^2} // \text{Simplify}$

Out[5]= 
$$\frac{2 p^4 \sigma^8}{(-2 b + c + d)^3}$$

Now all I need to do is prove that this quantity is positive.

Thank you for listening!

All the things Mathematica can do are  
documented here, with examples:

<https://reference.wolfram.com/language/>