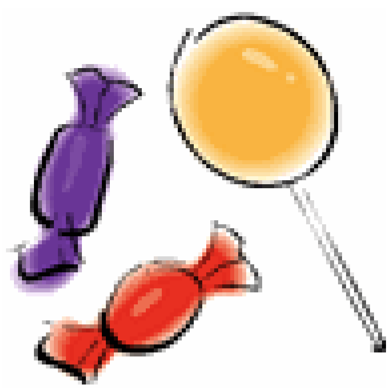
The Problem with the “~~Junk Food~~ Problem”

Snacks



Stephen Szydlik
University of Wisconsin Oshkosh
szydliks@uwosh.edu

Please contact Steve Szydlik at *szydliks@uwosh.edu* if you would like the Powerpoint version of this presentation. (This .pdf version does not allow proper representation of the animated GIFs in several of the slides.)

Themes

- *"….Problems that occur in real settings do not often arrive neatly packaged."*

  – *NCTM Standards 2000(p. 334)*

- *"I have yet to see any problem, however complicated, which, when you looked at it in the right way, did not become still more complicated."*

  – *Poul Anderson (1926-2001)*

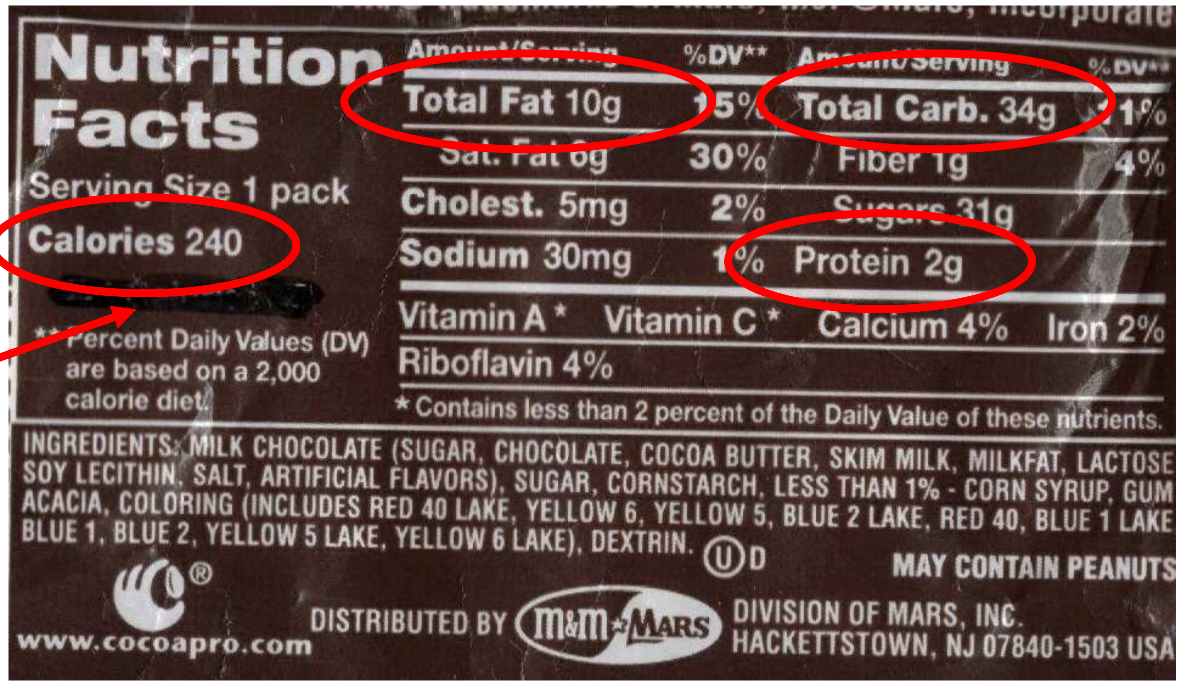
The Junk Food Problem

The calories in anything you eat come from three sources:

carbohydrates, fat, and protein.

Given the dietary information from some junk foods, find the number of calories that are present in a gram of carbohydrates, in a gram of fat and in a gram of protein.

Example:



(Solution blacked out.)

$$\Longrightarrow \quad 10f + 34c + 2p = 240$$

My (Original) Plan

- Give introductory linear algebra students some modeling experience.
- Introduce students to linear algebra applications, specifically solving linear systems using an augmented matrix.
- “Junk Food Problem” was to be just one of several applications to consider during the class period.
- Light discussion on error and the challenges of working with real data.

The Plan for Students

1. Assign variables (*f*, *c*, *p*).
2. Use snack wrappers to collect data and set up a 3x3 linear system.
3. Write the system as an augmented matrix, then row-reduce the matrix using technology.
4. Result: *f=9 cal/g*, *c=4 cal/g*, *p=4 cal/g*.

The Data

Item	Fat (g)	Carbs (g)	Protein (g)	Total Calories
m&m 's	10	34	2	240
Cheez Its	16	31	7	290
Cracker Jacks	1.5	21	2	100
Oreos	10	36	2	240
Twix	14	37	3	280
NutriGrain	3	27	2	140
Trail Mix	20	23	11	295
Planter's Peanuts	25	9	13	300

$$\Rightarrow \left[\begin{array}{ccc:c} 10 & 34 & 2 & 240 \\ 16 & 31 & 7 & 290 \\ 1.5 & 21 & 2 & 100 \\ 10 & 36 & 2 & 240 \\ 14 & 37 & 3 & 280 \\ 3 & 27 & 2 & 140 \\ 20 & 23 & 11 & 295 \\ 25 & 9 & 13 & 300 \end{array}\right]$$

Example

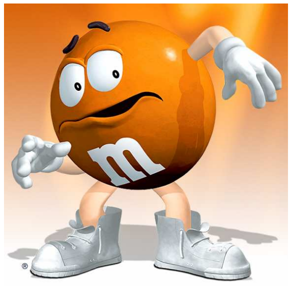
	fat	carb	prot	Cal.
CheezIts	**16**	**31**	**7**	**290**
Cracker Jacks	**1.5**	**21**	**2**	**100**
Twix	**14**	**37**	**3**	**280**

$$\Rightarrow \left[\begin{array}{ccc:c} 16 & 31 & 7 & 290 \\ 1.5 & 21 & 2 & 100 \\ 14 & 37 & 3 & 280 \end{array}\right]$$

$$\xrightarrow{\text{rref}}$$

$$\left[\begin{array}{ccc:c} 1 & 0 & 0 & 9.302 \\ 0 & 1 & 0 & 3.765 \\ 0 & 0 & 1 & 3.496 \end{array}\right] \Rightarrow$$

$$\boldsymbol{f \approx 9.302 \text{ cal/g}}$$
$$\boldsymbol{c \approx 3.765 \text{ cal/g}}$$
$$\boldsymbol{p \approx 3.496 \text{ cal/g}}$$



Problem!

	fat	carb	prot	Cal.
m&m's	10	34	2	240
Cracker Jacks	1.5	21	2	100
Twix	14	37	3	280

$$\Rightarrow \left[\begin{array}{ccc:c} 10 & 34 & 2 & 240 \\ 1.5 & 21 & 2 & 100 \\ 14 & 37 & 3 & 280 \end{array}\right]$$

$$\Big\downarrow \text{rref}$$

$$\Leftarrow \left[\begin{array}{ccc:c} 1 & 0 & 0 & 8.679 \\ 0 & 1 & 0 & 5.094 \\ 0 & 0 & 1 & -10.0 \end{array}\right]$$

$f \approx$ 8.679 cal/g
$c \approx$ 5.094 cal/g
$p \approx$ -10.0 cal/g

Measuring Error

We want to solve the linear system $Ax=b$.

Given an approximate solution x_0, we can calculate residual vector $b\text{-}Ax_0$ and measure the size $\|b\text{-}Ax_0\|$ to check the quality of the solution. We can measure the relative error by measuring

$$\frac{\|b - Ax_0\|}{\|b\|}.$$

Example: m&m's, Cracker Jacks, Twix

$$\text{Given} \quad A = \begin{bmatrix} 10 & 34 & 2 \\ 1.5 & 21 & 2 \\ 14 & 37 & 3 \end{bmatrix}, \quad x = \begin{bmatrix} f \\ c \\ p \end{bmatrix}, \quad b = \begin{bmatrix} 240 \\ 100 \\ 280 \end{bmatrix},$$

When $x_0 = \begin{bmatrix} 8.679 \\ 5.094 \\ -10.0 \end{bmatrix}$, $\|b - Ax_0\|$ and $\dfrac{\|b - Ax_0\|}{\|b\|}$. are essentially 0 (of course).

But when $x_0 = \begin{bmatrix} 9 \\ 4 \\ 4 \end{bmatrix}$, $\dfrac{\|b - Ax_0\|}{\|b\|} \approx 0.0265.$

The Point: Even though the data yield a horrible solution, the accepted solution is not a bad answer either.

Regression?

- Using all of the data yields an overdetermined system of 8 equations and 3 unknowns.
- Use regression for a least-squares solution:

Given $A = \begin{bmatrix} 10 & 34 & 2 \\ 16 & 31 & 7 \\ 1.5 & 21 & 2 \\ 10 & 36 & 2 \\ 14 & 37 & 3 \\ 3 & 27 & 2 \\ 20 & 23 & 11 \\ 25 & 9 & 13 \end{bmatrix}$ $x = \begin{bmatrix} f \\ c \\ p \end{bmatrix}$ $b = \begin{bmatrix} 240 \\ 290 \\ 100 \\ 140 \\ 280 \\ 140 \\ 295 \\ 300 \end{bmatrix}$

Solve $A^T A x = A^T b$ to find

***f* ≈ 7.398 cal/g, *c* ≈ 3.518 cal/g, *p* ≈ 6.690 cal/g**

Investigation

Go down a dimension and consider two different linear systems:

$$\begin{cases} 0.25x - y = -1 \\ 2x - y = 6 \end{cases} \quad \text{versus} \quad \begin{cases} 0.25x - y = -1 \\ 0.2x - y = -1.2 \end{cases}$$

Both systems have the solution (4,2).

But what if the systems arose from real data?

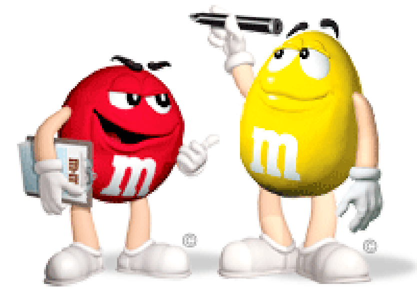
Change the constant terms slightly:

$$\begin{cases} 0.25x - y = -0.9 \\ 2x - y = 6 \end{cases} \qquad \begin{cases} 0.25x - y = -0.9 \\ 0.2x - y = -1.2 \end{cases}$$

Solution≈(3.9,1.9) **Solution (6,2.4)**

A very small change to the second system caused a HUGE change in the solution. WHY?

The Algebra



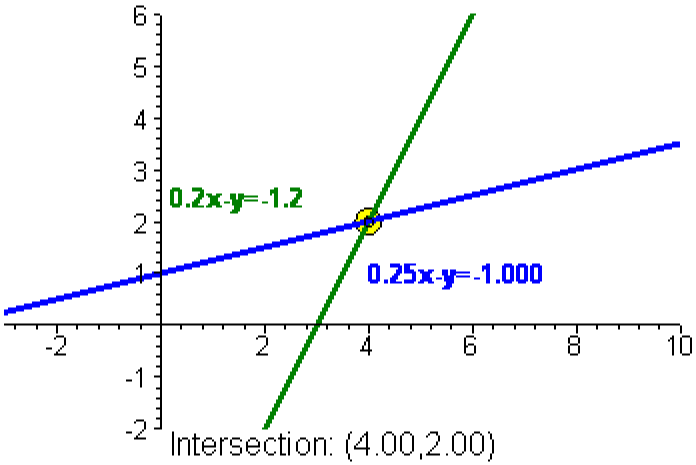
$$\begin{cases} 0.25x - y = -\cancel{1.0} \;\; \mathbf{-0.9} \\ 0.20x - y = -1.2 \end{cases}$$

$$0.05x = \cancel{0.2} \;\; \mathbf{0.3}$$

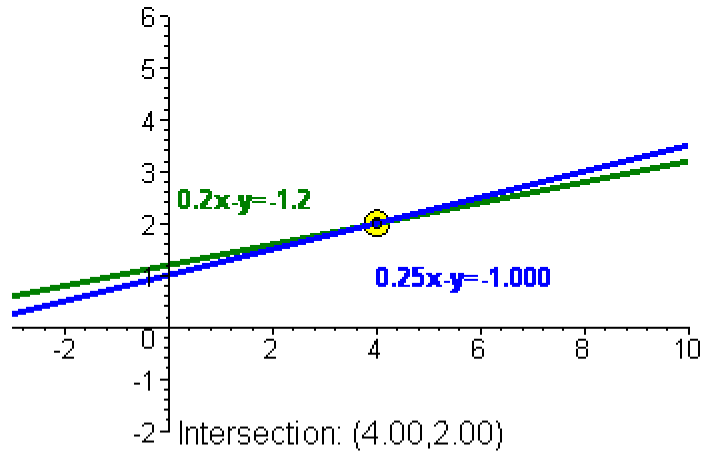
$$x = \cancel{4} \;\; \mathbf{6}$$

The left hand sides of the two equations are similar. This leads to an *x* with a very small coefficient after elimination of the y variable. So a small change in the right-hand constant causes a dramatic change in the solution when we divide by that coefficient.

The Geometry

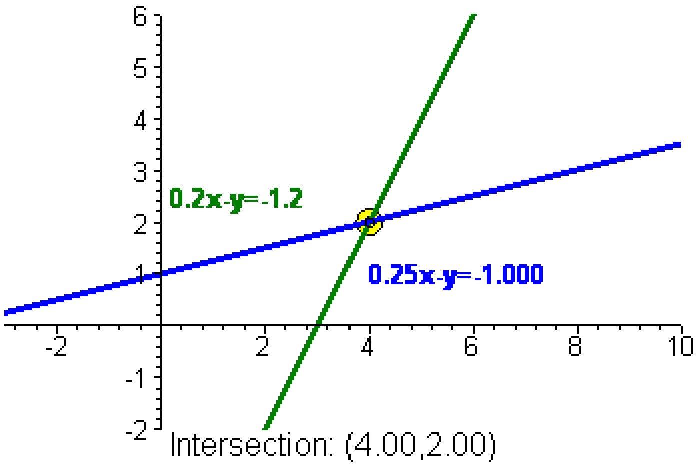


Well-conditioned
System

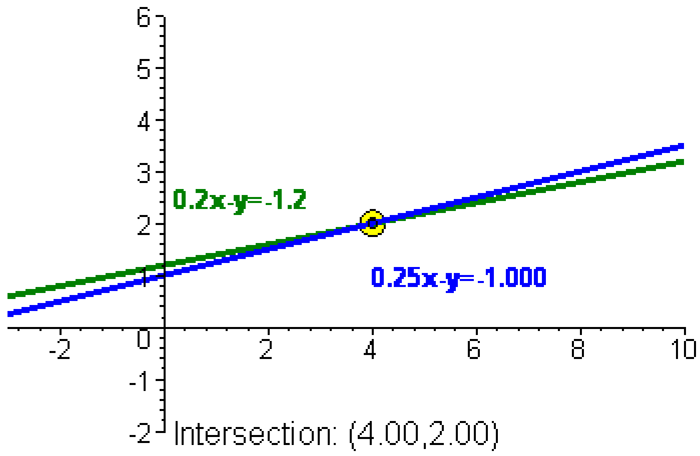


Ill-conditioned
System

The Geometry

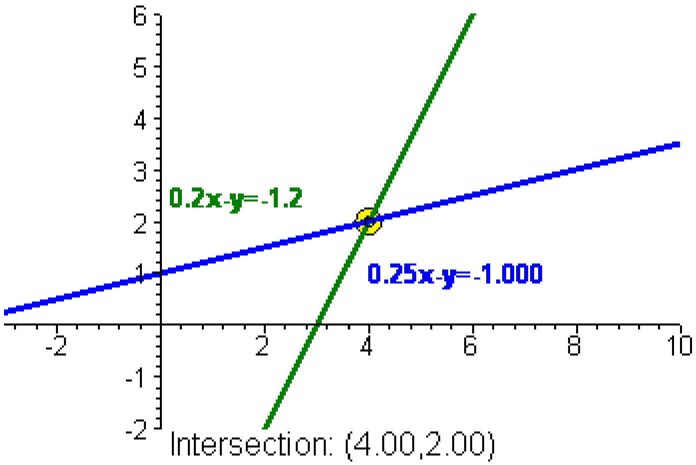


Well-conditioned System

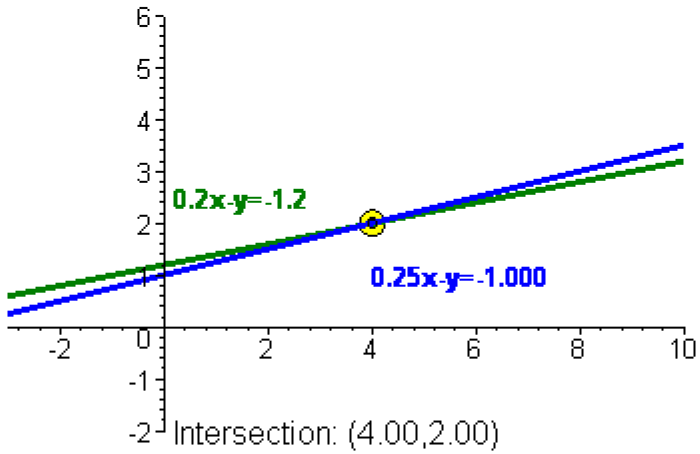


Ill-conditioned System

The Geometry

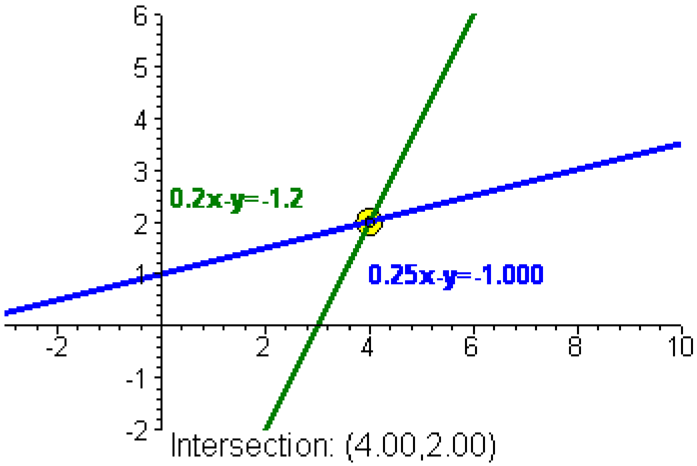


Well-conditioned
System

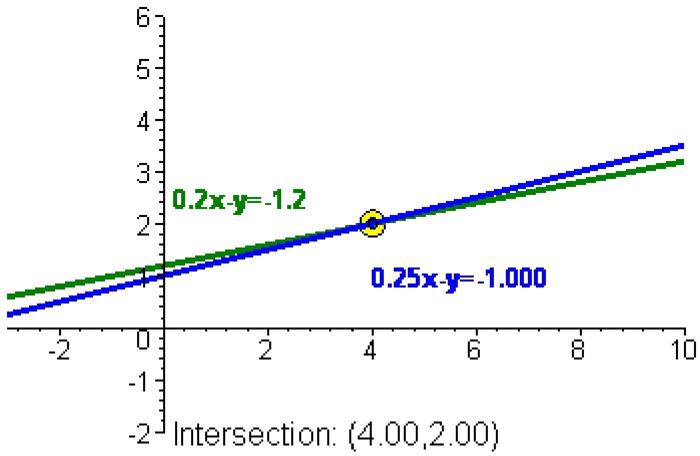


Ill-conditioned
System

The Geometry



Well-conditioned
System

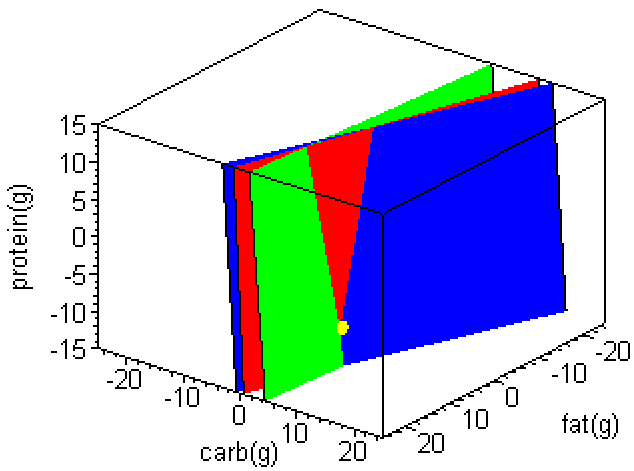


Ill-conditioned
System

Is this the Problem with the "Junk Food Problem?"

m&m's: $10f + 34c + 2p = 240$
Cracker Jacks: $1.5f + 21c + 2p = 100$
Twix: $14f + 37c + 3p = 280$

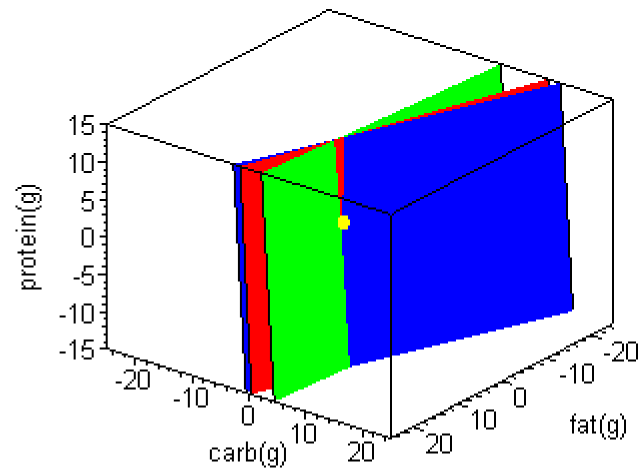
Intersection: (8.7, 5.1,-10.0)



"Actual Data?"

m&m's: $10.4f + 34.2c + 2.3p = 239.6$
Cracker Jacks: $1.3f + 20.6c + 1.55p = 100.3$
Twix: $13.6f + 36.6c + 2.7p = 279.6$

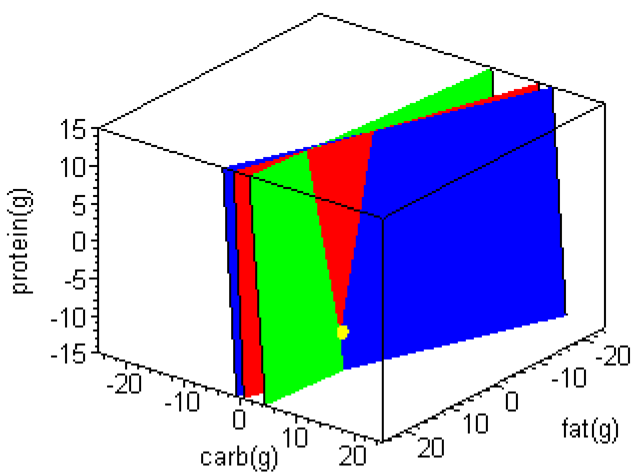
Intersection: (9.0, 4.0, 4.0)



Is this the Problem with the “Junk Food Problem?”

m&m’s: $10\,f + 34\,c + 2\,p = 240$
Cracker Jacks: $1.5\,f + 21\,c + 2\,p = 100$
Twix: $14\,f + 37\,c + 3\,p = 280$

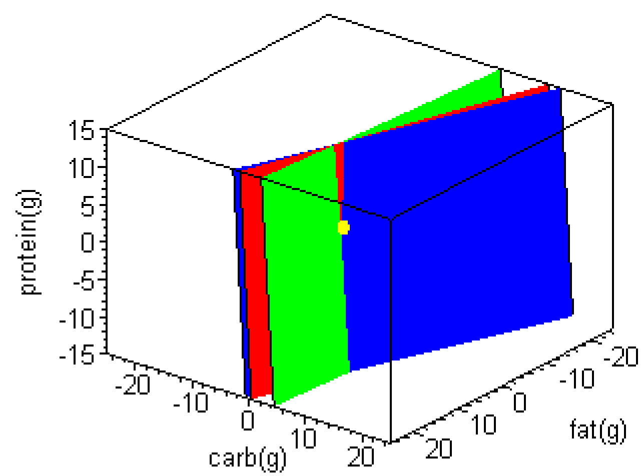
Intersection: (8.7, 5.1,-10.0)



“Actual Data?”

m&m’s: $10.4\,f + 34.2\,c + 2.3\,p = 239.6$
Cracker Jacks: $1.3\,f + 20.6\,c + 1.55\,p = 100.3$
Twix: $13.6\,f + 36.6\,c + 2.7\,p = 279.6$

Intersection: (9.0, 4.0, 4.0)



Resolution

The Problem: The data is too similar in its proportions of fat, carbs, and protein. In particular, most of the snacks are high fat, high carb, low protein.

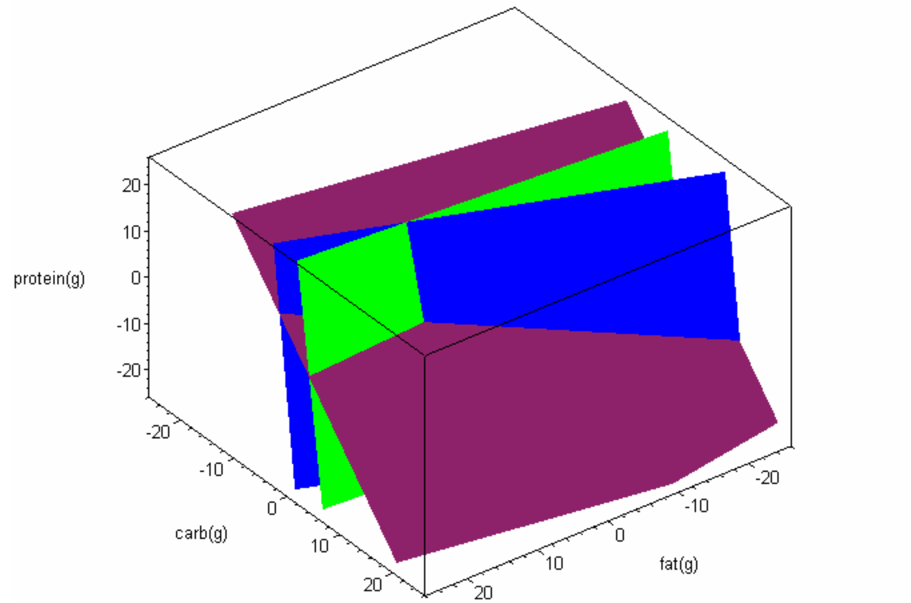
The Resolution: Include some more diverse data, using a higher protein snack.

Power Bar Protein Plus: $6f + 37c + 23p = 290$

Cracker Jacks: $1.5f + 21c + 2p = 100$

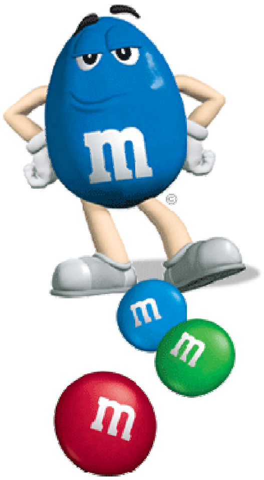
Twix: $14f + 37c + 3p = 280$

Solution ≈(9.34,3.69,4.23)



Final Thoughts

- Junk Food Problem as a springboard problem:
    1. *Condition number* of a matrix: measures how a small change in b affects the solutions to the linear system Ax=b.
    2. *Least Squares/Regression*
- Deeper discussion of solution error.



Stephen Szydlik
University of Wisconsin Oshkosh
szydliks@uwosh.edu