The Problem with the SNACK

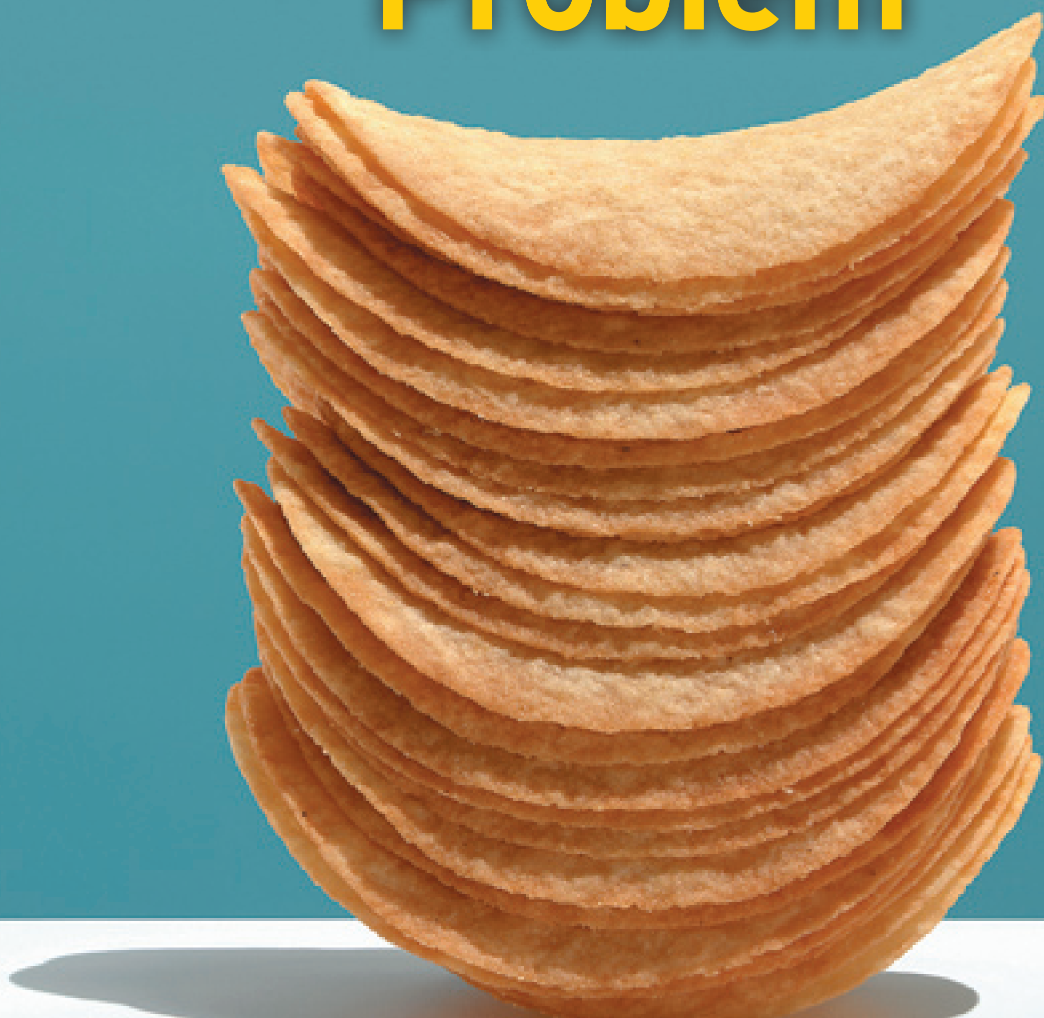
Solving systems of linear equations when using real data can present surprising challenges.

Stephen D. Szydlik

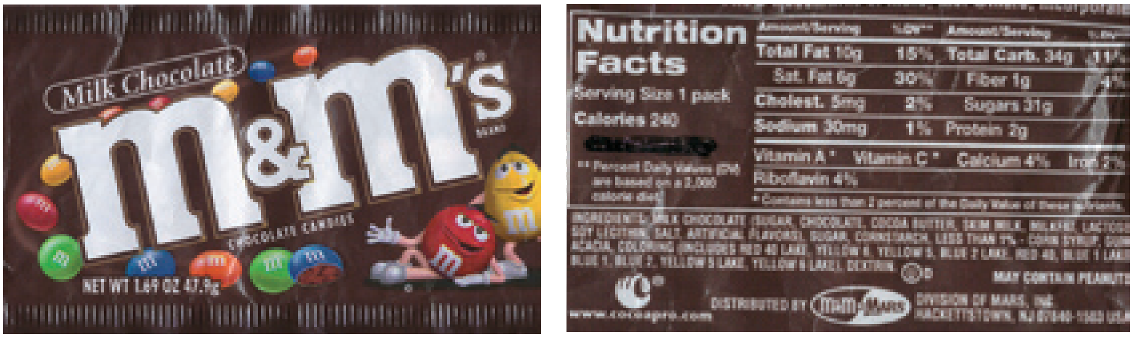
I have yet to see any problem, however complicated, which, when you looked at it in the right way, did not become still more complicated.
—Poul Anderson (1926–2001)

Early in their introductory college linear algebra class, my students learn to use a matrix to represent and solve systems of linear equations. Indeed, this simple idea is perhaps the foundational idea of the course. Nevertheless, I was delighted recently to find that the study of linear systems can lead to surprisingly deep mathematics. In this article, I describe a real-life class activity that seemed innocuous enough but presented unexpected challenges. Indeed, what began as an apparently straightforward exercise in setting up and solving a system of linear equations became an exploration of issues in numerical analysis, including error analysis and stability in such systems.

FOOD Problem



CHRIS STEIN/GETTY IMAGES



BEN SZYDLIK

Figure 1 Nutrition Information from an M&M's package. Many snack food wrappers provide the accepted solution to the Snack Food problem. Those data have been blacked out here.

THE ACTIVITY

Problems that occur in real settings do not often arrive neatly packaged.—NCTM 2000, p. 335

The Snack Food problem is one of those problems that do not arrive neatly packaged. The problem can be stated as follows:

> The calories in anything you eat come from three sources: carbohydrates, fat, and protein. Given the dietary information from some snack foods, find the number of calories present in a gram of carbohydrates, a gram of fat, and a gram of protein.

When my class began work on the Snack Food problem, we had finished a chapter on systems of linear equations. My students could represent such a system by using either an augmented coefficient matrix or the matrix form $AX = B$. Students had solved these types of systems both by hand and by using technology.

After arranging my students in groups of three or four, I distributed the problem statement and some snack food I had procured from the vending machine: candy, nuts, cookies, and chips. I gave each group one snack item, and the groups shared the information on the packages with one another in order to gather enough data to solve the problem. Each wrapper contained a wealth of nutritional information, some of which is helpful in answering the question. In particular, the "Nutrition Facts" label is required to provide, among other details, the total amount of fat, carbohydrates, and protein (all in grams) in a serving of the snack as well as the total number of calories in a serving. In principle, this information should have allowed my students to solve the Snack Food problem by setting up an appropriate system of linear equations.

Some words of caution before we begin our mathematical analysis: First, some snack labels contain too much information. Along with the number of grams of fat, some include the number of calories of fat, data that oversimplify the problem. Worse (in terms of the problem), some snack labels provide the entire solution (for example, fat: 9 cal/g; carbohydrate: 4 cal/g; protein: 4 cal/g). Black permanent marker, used judiciously, prevents a ruined problem (see **fig. 1**).

Second, some students may be aware of the accepted solution values. These students often have the most difficulty with the problem, because they need to argue how they know that those accepted values are correct; in some ways, having the "answer" presents a roadblock to their understanding of the solution.

Finally, it should be noted that the accepted values for the number of calories in a gram of fat (e.g., 9), a gram of carbohydrate (e.g., 4), and a gram of protein (e.g., 4), respectively, are approximations. The values, called the Atwater factors, are found by measuring the heat of combustion for the different substances. Typically, the values are averages, because different foods yield slightly different heat-of-combustion results. For example, eggs contain slightly more calories per gram of protein than do meats. For a more complete discussion, see Davidson and others (1973, pp. 8–10).

THE SOLUTION AND THE SURPRISE

Now, with a sense of the problem and with the data in front of them, the students got to work. The more accomplished students immediately saw how to attack the problem, and even the less sophisticated students needed little prodding to find an effective approach: Assign variables (e.g., f, c, and p) for the respective calories in a gram of fat, a gram of carbohydrate, or a gram of protein and use the information on the snack food wrappers to identify relationships among the unknowns.

I had purchased eight snacks from the vending machine; the relevant nutritional information is found in **table 1**. My students knew that they did not need all eight snacks to solve the problem, so they chose some subset. For example, Will's group collected data from Cheez-It®, Cracker Jack®, and Twix®. This group's setup looked something like that in **table 2**.

Solving the corresponding system of linear equations by hand is not difficult, but it is tedious. The matrix capabilities of a graphing calculator can eliminate the drudgery, allowing for more emphasis on analysis. As my students had learned, the coefficients of the unknowns in the system can be placed in an augmented matrix that can be row reduced. For example, Will's group's work had the following form:

$$\begin{pmatrix} 16 & 31 & 7 & 290 \\ 1.5 & 21 & 2 & 100 \\ 14 & 37 & 3 & 280 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & 9.302 \\ 0 & 1 & 0 & 3.765 \\ 0 & 0 & 1 & 3.496 \end{pmatrix}$$

This group thus arrived at the approximate solution $f \approx 9.302$ cal/g, $p \approx 3.765$ cal/g, and $c \approx 3.496$ cal/g. This solution is a reasonable approximation of the accepted values. Other groups solved the problem in a similar way, although using different combinations of the available snacks. I had expected that these other groups would provide slightly different approximate solutions, an outcome that would lead to a relatively light discussion of the challenges of dealing with real data.

However, as the groups began reporting their solutions, it became apparent that something was wrong. Laurie called me over to check her group's solution; these students had gathered their data from the M&M's®, Cheez-It, and Nutri•Grain® snacks. They solved their system

$$\begin{aligned} 10f + 34c + 2p &= 240 && \text{M\&M's} \\ 16f + 31c + 7p &= 290 && \text{Cheez-It} \\ 3f + 27c + 2p &= 140 && \text{Nutri•Grain} \end{aligned}$$

to find $f \approx 10.248$, $c \approx 4.037$, and $p \approx 0.124$.

Erin's group used M&M's, Cracker Jack, and Twix to solve the system

$$\begin{aligned} 10f + 34c + 2p &= 240 && \text{M\&M's} \\ 1.5f + 21c + 2p &= 100 && \text{Cracker Jack} \\ 14f + 37c + 3p &= 280 && \text{Twix} \end{aligned}$$

and, even more inexplicably, obtained $f \approx 8.679$, $c \approx 5.094$, and $p \approx -10$. Other groups found similarly problematic answers.

ERROR ANALYSIS

I asked my students to check their answers. Perhaps the simplest check is to substitute the values obtained for f, c, and p into the equations for the system and review that these values do make the equations simultaneously true.

Table 1

Fat, Carbohydrates, Protein, and Total Calories in a Single Serving of Selected Snacks from the Vending Machine

Item	Fat (g)	Carbs (g)	Protein (g)	Calories
M&M's®	10	34	2	240
Cheez-It®	16	31	7	290
Cracker Jack®	1.5	21	2	100
Oreo®	10	36	2	240
Twix®	14	37	3	280
Nutri•Grain®	3	27	2	140
Trail mix	20	23	11	295
Planters® peanuts	25	9	13	300

Table 2

Will's Group's System of Equations

Item	*f*	*c*	*p*	Cal.		
Cheez-It	16	31	7	290	→	$16f + 31c + 7p = 290$
Cracker Jack	1.5	21	2	100	→	$1.5f + 21c + 2p = 100$
Twix	14	37	3	280	→	$14f + 37c + 3p = 280$

A similar but slightly more sophisticated approach to checking an answer involves matrix multiplication and the matrix form of a linear system. Students familiar with these ideas can express their systems in the form $AX = B$, where A is the coefficient matrix for the linear system, X is the solution vector $[f, c, p]^T$, and B is the vector of the constants on the right side of the equations. So, for example, Erin's group could express its linear system in the matrix form $AX = B$ or as shown below:

$$\begin{bmatrix} 10 & 34 & 2 \\ 1.5 & 21 & 2 \\ 14 & 37 & 3 \end{bmatrix} \begin{bmatrix} f \\ c \\ p \end{bmatrix} = \begin{bmatrix} 240 \\ 100 \\ 280 \end{bmatrix}$$

Then, checking the solutions amounted to checking that the values for f, c, and p made the matrix equation true. The advantage of this approach is that when the two sides of the matrix equation do not agree, the difference between those two sides provides a means of quantifying the "correctness" of the solution.

Using the solution that had been rounded at the thousandths place, Erin's group used this approach and found that

$$AX = \begin{bmatrix} 10 & 34 & 2 \\ 1.5 & 21 & 2 \\ 14 & 37 & 3 \end{bmatrix} \begin{bmatrix} 8.679 \\ 5.094 \\ -10 \end{bmatrix} = \begin{bmatrix} 239.986 \\ 99.9925 \\ 279.984 \end{bmatrix},$$

compared with

$$B = \begin{bmatrix} 240 \\ 100 \\ 280 \end{bmatrix}.$$

Thus, the group's *residual,* or *error, vector* was

$$E = B - AX = \begin{bmatrix} .014 \\ .0075 \\ .016 \end{bmatrix}.$$

The *magnitude* of the error vector provides a single measure of the size of the error:

$$\|E\| = \sqrt{0.014^2 + 0.0075^2 + 0.016^2} \approx 0.02254$$

Although the group's residual error is small, we can get a better sense of the quality of the solution by comparing the magnitude of the error with the magnitude of vector B:

$$\frac{\|B - AX\|}{\|B\|} = \frac{\|E\|}{\|B\|} = \frac{\|E\|}{\sqrt{240^2 + 100^2 + 280^2}}$$

$$\approx \frac{0.02254}{382.099} \approx 0.000059,$$

or about 0.0059%. This *relative error* tells us that for Erin's group's solution X, the matrix product AX is within about 0.0059% of B. Her group's answer is a correct solution to the system of equations, although it does not make sense as a solution to the problem. (The only reason the residual and relative errors are not identical—zero—is because the group rounded its solution to the nearest thousandth.)

What went wrong? How could the students get so many apparently correct solutions that were so different from one another? And why were these solutions so far from the accepted solution? The difficulty lies in the theoretical construct (in particular, the linear relationship among the three calorie sources), in the data, or in both.

One simple way to check both the theoretical construct and the data is to find the residual error in the accepted solution $X_0 = [9, 4, 4]^T$. This approach measures how closely the accepted solution fits the linear system. Using Erin's group's data, for example, gives a residual error of

$$\|B - AX_0\| = \left\| \begin{bmatrix} 240 \\ 100 \\ 280 \end{bmatrix} - \begin{bmatrix} 10 & 34 & 2 \\ 1.5 & 21 & 2 \\ 14 & 37 & 3 \end{bmatrix} \bullet \begin{bmatrix} 9 \\ 4 \\ 4 \end{bmatrix} \right\|$$

$$= \left\| \begin{bmatrix} 240 \\ 100 \\ 280 \end{bmatrix} - \begin{bmatrix} 234 \\ 105.5 \\ 286 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 6 \\ -5.5 \\ -6 \end{bmatrix} \right\| \approx 10.11$$

and a relative residual error of

$$\frac{\|B - AX_0\|}{\|B\|} \approx 0.0265.$$

By this measure, then, AX_0 is within 2.7% of B in this case. So although the error is nonnegligible, it remains small in relation to the size of B. When other groups used the accepted values in their systems, they obtained residual errors of similar magnitude. In short, the accepted values (in cal/g) of $f = 9$, $c = 4$, and $p = 4$ do provide reasonable approximate solutions to the linear systems that the different groups set up. The source of our problem, then, is likely not the theoretical framework. The assumption that the total calories in our snack foods are a linear combination of the grams of fat, carbohydrate, and protein appears sound. In addition, this result tells us that the snack food data also appear reliable.

INVESTIGATION

Understand simple things deeply. —Burger and Starbird 2000, p. 25

As one important problem-solving strategy, NCTM's *Principles and Standards for School Mathematics* (2000) recommends reducing a problem to one that is simpler, and by doing so my students and I were finally able to identify the source of confusion in the Snack Food problem. We explored an analogous problem that involved two unknowns instead of three.

Consider the following two linear systems of two equations and two unknowns:

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

Although the systems are entirely concocted, we may think of the equations as arising from some physical application.

At first glance, the systems appear quite similar. Indeed, the first equations in both systems are identical, and the systems share the common solution $x = 4$, $y = 2$. However, the two systems are substantially different in their *stability*. If, indeed, the coefficients for the linear systems were collected as real data, there would necessarily be some uncertainty about the accuracy of their values. How certain can we then be of our solution? For example, suppose that in both systems, the constant on the right side of the first equation actually has the value of –0.9, rather than –1 (a 10% difference). So the systems would be more correctly expressed as follows:

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

With this slight perturbation, we may reasonably expect the solutions to change slightly but not substantially. When we solve the first system, we find that $x = 138/35 \approx 3.943$ and $y = 66/35 \approx 1.886$. These values correspond reasonably with the original solution $x = 4$, $y = 2$ (there is a change of approximately 1.4% in the x-coordinate and a change of approximately 5.7% in the y-coordinate). However, when we solve the perturbed second system, we find that $x = 6$, $y = 2.4$. This is a dramatic change; the effect on the x-coordinate of the solution is a 50 percent change (from $x = 4$ to $x = 6$). So a small change in both linear systems causes a fairly minor change in the solution to the first system but a significant change in the solution to the second system. The first system displays a robustness that the second system lacks.

What causes this dramatic difference in the stability of the two systems? A look at the graphs of the two linear systems is illustrative (see **figs. 2** and **3**). In both figures, the blue line represents the graph of the equation $0.25x - y = -1$, while the green line represents the graph of the second equation. The red dashed line gives the graph of the perturbed equation $0.25x - y = -0.9$. In each case, changing the constant from –1 to –0.9 corresponds graphically to changing the y-intercept of the blue line, in effect causing the blue line to slide down slightly. Now, in the first system, the slopes of the green and blue lines are significantly different, and that small downward slide by the blue line causes little change in the intersection point of the two lines. In the second system, in contrast, the lines have slopes that are extremely close. So changing the y-intercept of the blue line affects the intersection point of the two lines significantly. This effect can be demonstrated visually by using a dynamic geometry package such as The Geometer's Sketchpad® or simply with a pair of metersticks.

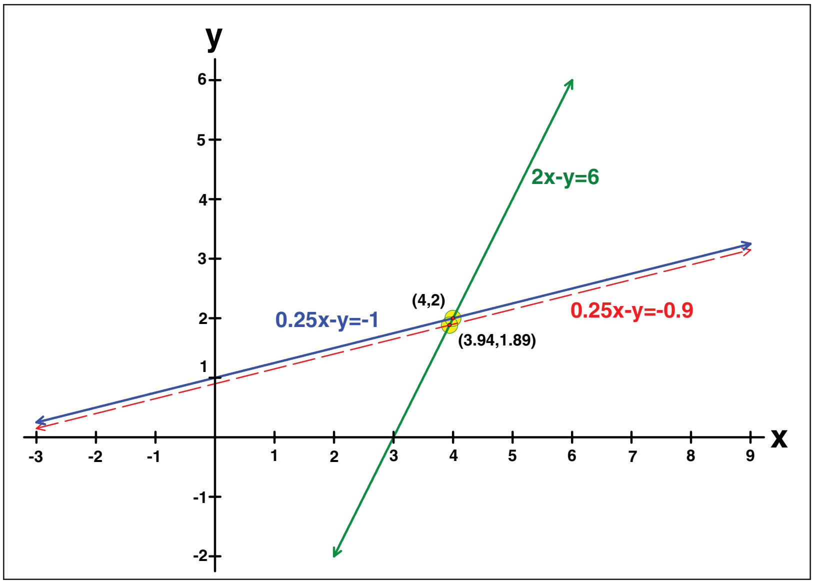


Fig. 2 The graph of a stable linear system in two dimensions. Perturbing the blue line slightly (to the red line) yields a correspondingly small change in the solution to the system.

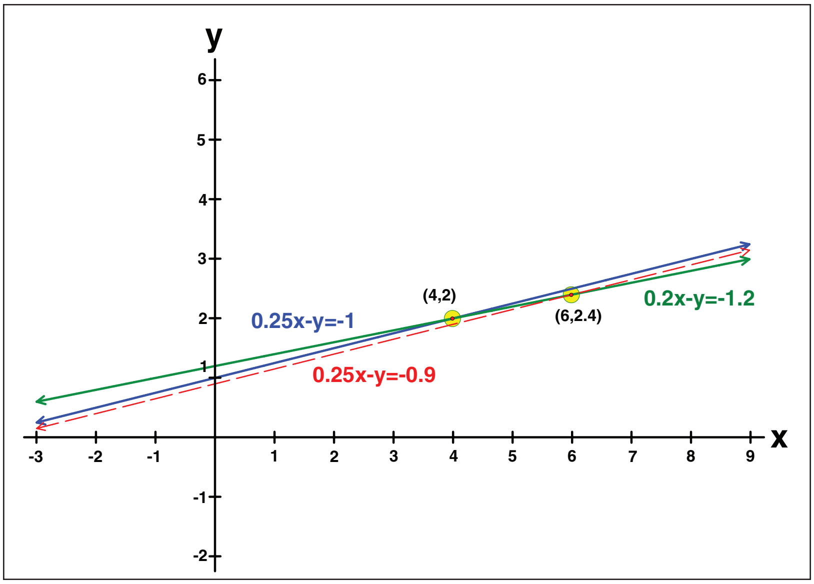


Fig. 3 The graph of an unstable linear system in two dimensions. Perturbing the blue line slightly (to the red line) causes a substantial change in the solution to the system.

We can see this effect in the algebra of the solution to the second system as well. Using simple elimination, we can subtract the second equation from the first in the unperturbed system to eliminate the y variable:

$$\begin{cases} 0.25x - y = -1 \\ 0.2x - y = -1.2 \end{cases} \rightarrow 0.05x = 0.2 \rightarrow x = 4$$

Because of the similarity of the left sides of the equations in the system, the coefficient of x after the elimination of y is a small positive number. As a consequence, dividing by that coefficient makes the right side substantially larger. When we perturb the right constant in the first equation to –0.9, we obtain the single-variable equation $0.05x = 0.3$. It is a small change on the right side, *but that change of 0.1 grows twentyfold when we divide by 0.05 to solve for x*. This fact causes the dramatic change in the solution.

RESOLUTION

Might this be the problem with the Snack Food problem? Let's return to Erin's group's system of equations:

$$\begin{cases} 10f + 34c + 2p = 240 \\ 1.5f + 21c + 2p = 100 \\ 14f + 37c + 3p = 280 \end{cases},$$

which yields $f \approx 8.679$, $c \approx 5.094$, and $p = -10$. How great is the effect on the solution if we change one or more of the coefficients in her group's system? For example, suppose that in the first piece of data, the M&M's actually have 10.1 grams of fat rather than 10 (a relative difference of 1%). This is not an unreasonable possibility; the data in **table 1** appear to be rounded to the nearest whole number in most cases. With this change, Erin's group's system becomes

$$\begin{cases} 10.1f + 34c + 2p = 240 \\ 1.5f + 21c + 2p = 100 \\ 14f + 37c + 3p = 280 \end{cases},$$

with solution $f \approx 8.725$, $c \approx 4.998$, and $p \approx -9.018$.

Although the approximate values of f and c remain similar under this minor perturbation of 0.1, the value of p changes by almost 10 percent. The system demonstrates a large measure of instability. We can demonstrate this idea even more dramatically by perturbing other coefficients of Erin's group's system but still keeping them within the limits of rounding.

For example, suppose that the actual number of grams of fat, carbohydrate, and protein in Erin's group's samples are given as in **table 3**. Although these values have been concocted, they are consistent with the data given in **table 1** up to rounding. Solving the corresponding linear system in this case yields $f = 9$, $c = 4$, and $p = 4$. In other words, Erin's group's system is sensitive enough to perturbation that its entire deviation from the accepted solution could be explained by round-off error in the data.

Table 3

Concocted Values Used to Investigate Perturbation

Item	*f*	*c*	*p*	Cal.
M&M's®	10.4	34.2	2.3	239.6
Cracker Jack	1.3	20.6	1.55	100.3
Twix	13.6	36.6	2.7	279.6

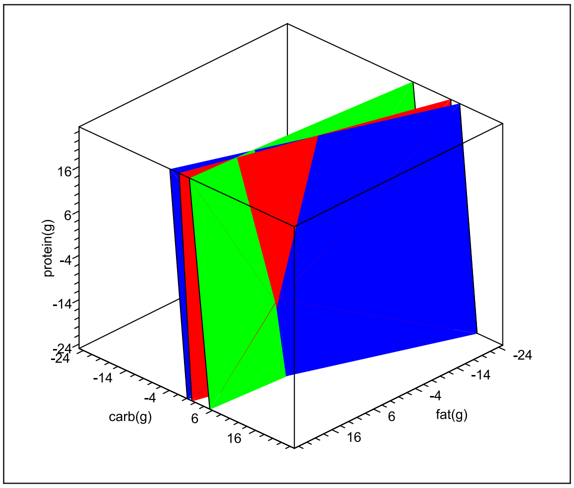


Fig. 4 The graph of an unstable linear system in three dimensions

As in the two variable examples, the instability in the linear system can be seen visually. The graph of each of Erin's group's equations can be represented as a plane in three dimensions (with f, c, and p as axes). We color the planes corresponding to the M&M's, Cracker Jack, and Twix data red, green, and blue, respectively (see **fig. 4**).

The intersection of the three planes is the solution point with approximate coordinates (8.679, 5.094, –10.000). The three planes are close to parallel and also nearly vertical. Small changes in the vertical intercept of any of these planes, therefore, has only a small effect on the solution, but changing the "tilt" of the planes only slightly can have a profound effect, as we have seen.

The issue with the Snack Food problem, then, is indeed with the data. However, the difficulty is not inaccuracy of the individual snack data; rather, the problem occurs with the data set as a whole—the data are not diverse enough. Many snacks available from vending machines tend to be high in fat and carbohydrates and low in protein *and in reasonably similar proportions*. Systems derived from these data look similar, in the sense described above. In these cases, small perturbations to the system caused by rounding the data cause distressing results, such as the one my class experienced.

The resolution was clear to my students: We needed data that were more related to nutrition. The fix we chose was to replace one of the high-fat, high-calorie, low-protein snacks with a high-protein snack, namely a PowerBar® ProteinPlus™. This energy bar contains 6 grams of fat, 37 grams of carbohydrates, and an impressive 23 grams of protein. If we replace the M&M's data with the PowerBar data, we obtain the following linear system:

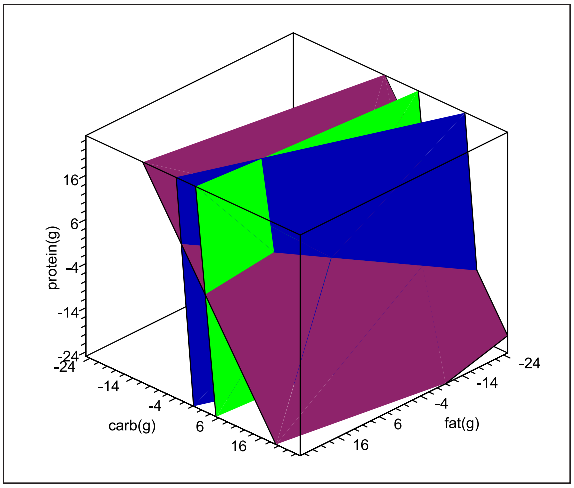


Fig. 5 The graph of a stable linear system in three dimensions

$$\begin{cases} 6f + 37c + 23p = 290 \\ 1.5f + 21c + 2p = 100 \\ 14f + 37c + 3p = 280 \end{cases},$$

with solution $f \approx 9.336$, $c \approx 3.692$, and $p \approx 4.234$. This solution is quite close to the accepted values. Moreover, the system is now significantly more stable: A slight perturbation of the coefficients causes a correspondingly small change in the solution.

Once again, this result can be observed graphically. As before, we color the planes corresponding to the Cracker Jack and Twix data green and blue, respectively, and we replace the plane corresponding to the M&M's data with the plane of the PowerBar ProteinPlus data, coloring it maroon (see **fig. 5**).

The mathematics of the Snack Food problem happens to parallel its nutritional aspects. Just as our mathematical troubles in the problem arose from relying too heavily on high-fat, high-carbohydrate, low-protein snacks, so too can one get into trouble nutritionally from relying too heavily on snack foods in one's diet. In either case, solving the Snack Food problem requires identifying more nutritional snacks. Snack diversity is desirable, both nutritionally and mathematically!

FURTHER STUDY

Stability is a characteristic of the coefficient matrix of the system and can be measured by finding the *condition number* of the matrix (see Atkinson [1978], p. 457; see Kalman [1996] for an excellent discussion of conditioning and its relationship to least-squares analysis). Although the discussion is beyond the scope of this article, the condition number allows us to quantify the amount of confidence we can have in our solution. The Snack Food problem reveals the significance of stability and why conditioning is an important aspect in the study of linear systems.

CONCLUSION

Despite my disconcerting classroom encounter with the Snack Food problem, I continue to use it in my linear algebra course. However, I no longer view it as a routine exercise. Rather, I use this problem precisely for the complexity it affords. This complexity offers my students a valuable mathematical experience and provides avenues leading to deeper mathematical discussions of error analysis, stability of linear systems, and (later in the semester) linear least-squares analysis.

NCTM's *Principles and Standards for School Mathematics* notes that approaching mathematical content through problem solving "reveals mathematics as a sense-making discipline rather than one in which rules for working exercises are given by the teacher to be memorized and used by students" (p. 334). For me, the Snack Food problem embodies this approach in a most powerful way.

BIBLIOGRAPHY

Anderson, Poul. Quoted in William Thorpe, "Reduction v. Organicism," *New Scientist* 25, no. 43 (September 1969): 638.

Atkinson, Kendall. *An Introduction to Numerical Analysis*. New York: Wiley, 1978.

Bretscher, Otto. *Linear Algebra with Applications*. Upper Saddle River, NJ: Pearson Prentice Hall, 2005.

Burger, Edward B., and Michael Starbird. *The Heart of Mathematics: An Invitation to Effective Thinking*. Emeryville, CA: Key College Publishing, 2000.

Davidson, Sir Stanley, R. Passmore, and J. F. Brock. *Human Nutrition and Dietetics*. Churchill Livingstone: Edinburgh and London, 1973.

Kalman, Dan. "A Singularly Valuable Decomposition: The SVD of a Matrix." *College Mathematics Journal* 27, no. 1 (January 1996): 2–25.

National Council of Teachers of Mathematics (NCTM). *Principles and Standards for School Mathematics*. Reston, VA: NCTM, 2000.

STEPHEN D. SZYDLIK, szydliks@uwosh.edu, teaches mathematics at the University of Wisconsin–Oshkosh. He is always looking for new problems that engage students in the classroom. Although he is an avid long-distance runner and triathlete, his doctor recently told him that he needs to curtail his intake of snack foods.