Street-Fighting Mathematics

The Art of Educated Guessing and Opportunistic Problem Solving

Sanjoy Mahajan

Foreword by Carver A. Mead

The MIT Press
Cambridge, Massachusetts
London, England

For information about special quantity discounts, please email
special_sales@mitpress.mit.edu

Typeset in Palatino and Euler by the author using ConTEXt and PDFTEX

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Mahajan, Sanjoy, 1969–
Street-fighting mathematics : the art of educated guessing and opportunistic problem solving / Sanjoy Mahajan ; foreword by Carver A. Mead.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-262-51429-3 (pbk. : alk. paper) 1. Problem solving. 2. Hypothesis. 3. Estimation theory. I. Title.
QA63.M34 2010
510—dc22

2009028867

Printed and bound in the United States of America

10 9 8 7 6 5 4 3 2 1

1 Dimensions

Our first street-fighting tool is dimensional analysis or, when abbreviated, dimensions. To show its diversity of application, the tool is introduced with an economics example and sharpened on examples from Newtonian mechanics and integral calculus.

1.1 Economics: The power of multinational corporations

Critics of globalization often make the following comparison [25] to prove the excessive power of multinational corporations:

> In Nigeria, a relatively economically strong country, the GDP [gross domestic product] is $99 billion. The net worth of Exxon is $119 billion. "When multinationals have a net worth higher than the GDP of the country in which they operate, what kind of power relationship are we talking about?" asks Laura Morosini.

Before continuing, explore the following question:

▶ *What is the most egregious fault in the comparison between Exxon and Nigeria?*

The field is competitive, but one fault stands out. It becomes evident after unpacking the meaning of GDP. A GDP of $99 billion is shorthand for a monetary flow of $99 billion per year. A year, which is the time for the earth to travel around the sun, is an astronomical phenomenon that has been arbitrarily chosen for measuring a social phenomenon—namely, monetary flow.

Suppose instead that economists had chosen the decade as the unit of time for measuring GDP. Then Nigeria's GDP (assuming the flow remains steady from year to year) would be roughly $1 trillion per decade and be reported as $1 trillion. Now Nigeria towers over Exxon, whose puny assets are a mere one-tenth of Nigeria's GDP. To deduce the opposite conclusion, suppose the week were the unit of time for measuring GDP. Nigeria's GDP becomes $2 billion per week, reported as $2 billion. Now puny Nigeria stands helpless before the mighty Exxon, 50-fold larger than Nigeria.

A valid economic argument cannot reach a conclusion that depends on the astronomical phenomenon chosen to measure time. The mistake lies in comparing incomparable quantities. Net worth is an amount: It has dimensions of money and is typically measured in units of dollars. GDP, however, is a flow or rate: It has dimensions of money per time and typical units of dollars per year. (A dimension is general and independent of the system of measurement, whereas the unit is how that dimension is measured in a particular system.) Comparing net worth to GDP compares a monetary amount to a monetary flow. Because their dimensions differ, the comparison is a category mistake [39] and is therefore guaranteed to generate nonsense.

Problem 1.1 Units or dimensions?

Are meters, kilograms, and seconds units or dimensions? What about energy, charge, power, and force?

A similarly flawed comparison is length per time (speed) versus length: "I walk $1.5\,\mathrm{m\,s^{-1}}$—much smaller than the Empire State building in New York, which is 300 m high." It is nonsense. To produce the opposite but still nonsense conclusion, measure time in hours: "I walk 5400 m/hr—much larger than the Empire State building, which is 300 m high."

I often see comparisons of corporate and national power similar to our Nigeria–Exxon example. I once wrote to one author explaining that I sympathized with his conclusion but that his argument contained a fatal dimensional mistake. He replied that I had made an interesting point but that the numerical comparison showing the country's weakness was stronger as he had written it, so he was leaving it unchanged!

A dimensionally valid comparison would compare like with like: either Nigeria's GDP with Exxon's revenues, or Exxon's net worth with Nigeria's net worth. Because net worths of countries are not often tabulated, whereas corporate revenues are widely available, try comparing Exxon's annual revenues with Nigeria's GDP. By 2006, Exxon had become Exxon Mobil with annual revenues of roughly $350 billion—almost twice Nigeria's 2006 GDP of $200 billion. This valid comparison is stronger than the flawed one, so retaining the flawed comparison was not even expedient!

That compared quantities must have identical dimensions is a necessary condition for making valid comparisons, but it is not sufficient. A costly illustration is the 1999 Mars Climate Orbiter (MCO), which crashed into the surface of Mars rather than slipping into orbit around it. The cause, according to the Mishap Investigation Board (MIB), was a mismatch between English and metric units [26, p. 6]:

> The MCO MIB has determined that the root cause for the loss of the MCO spacecraft was the failure to use metric units in the coding of a ground software file, Small Forces, used in trajectory models. Specifically, thruster performance data in English units instead of metric units was used in the software application code titled SM_FORCES (small forces). A file called Angular Momentum Desaturation (AMD) contained the output data from the SM_FORCES software. The data in the AMD file was required to be in metric units per existing software interface documentation, and the trajectory modelers assumed the data was provided in metric units per the requirements.

Make sure to mind your dimensions and units.

Problem 1.2 Finding bad comparisons

Look for everyday comparisons—for example, on the news, in the newspaper, or on the Internet—that are dimensionally faulty.

1.2 Newtonian mechanics: Free fall

Dimensions are useful not just to debunk incorrect arguments but also to generate correct ones. To do so, the quantities in a problem need to have dimensions. As a contrary example showing what not to do, here is how many calculus textbooks introduce a classic problem in motion:

> A ball initially at rest falls from a height of h **feet** and hits the ground at a speed of v **feet per second**. Find v assuming a gravitational acceleration of g **feet per second squared** and neglecting air resistance.

The units such as feet or feet per second are highlighted in boldface because their inclusion is so frequent as to otherwise escape notice, and their inclusion creates a significant problem. Because the height is h feet, the variable h does not contain the units of height: h is therefore dimensionless. (For h to have dimensions, the problem would instead state simply that the ball falls from a height h; then the dimension of length would belong to h.) A similar explicit specification of units means that the variables g and v are also dimensionless. Because g, h, and v are dimensionless, any comparison of v with quantities derived from g and h is a comparison between dimensionless quantities. It is therefore always dimensionally valid, so dimensional analysis cannot help us guess the impact speed.

Giving up the valuable tool of dimensions is like fighting with one hand tied behind our back. Thereby constrained, we must instead solve the following differential equation with initial conditions:

$$\frac{d^2y}{dt^2} = -g, \text{with } y(0) = h \text{ and } dy/dt = 0 \text{ at } t = 0, \tag{1.1}$$

where $y(t)$ is the ball's height, dy/dt is the ball's velocity, and g is the gravitational acceleration.

Problem 1.3 Calculus solution

Use calculus to show that the free-fall differential equation $d^2y/dt^2 = -g$ with initial conditions $y(0) = h$ and $dy/dt = 0$ at $t = 0$ has the following solution:

$$\frac{dy}{dt} = -gt \qquad \text{and} \qquad y = -\frac{1}{2}gt^2 + h. \tag{1.2}$$

▶ *Using the solutions for the ball's position and velocity in Problem 1.3, what is the impact speed?*

When $y(t) = 0$, the ball meets the ground. Thus the impact time t_0 is $\sqrt{2h/g}$. The impact velocity is $-gt_0$ or $-\sqrt{2gh}$. Therefore the impact speed (the unsigned velocity) is $\sqrt{2gh}$.

This analysis invites several algebra mistakes: forgetting to take a square root when solving for t_0, or dividing rather than multiplying by g when finding the impact velocity. Practice—in other words, making and correcting many mistakes—reduces their prevalence in simple problems, but complex problems with many steps remain minefields. We would like less error-prone methods.

One robust alternative is the method of dimensional analysis. But this tool requires that at least one quantity among v, g, and h have dimensions. Otherwise, every candidate impact speed, no matter how absurd, equates dimensionless quantities and therefore has valid dimensions.

Therefore, let's restate the free-fall problem so that the quantities retain their dimensions:

> A ball initially at rest falls from a height h and hits the ground at speed v. Find v assuming a gravitational acceleration g and neglecting air resistance.

The restatement is, first, shorter and crisper than the original phrasing:

> A ball initially at rest falls from a height of h feet and hits the ground at a speed of v feet per second. Find v assuming a gravitational acceleration of g feet per second squared and neglecting air resistance.

Second, the restatement is more general. It makes no assumption about the system of units, so it is useful even if meters, cubits, or furlongs are the unit of length. Most importantly, the restatement gives dimensions to h, g, and v. Their dimensions will almost uniquely determine the impact speed—without our needing to solve a differential equation.

The dimensions of height h are simply length or, for short, L. The dimensions of gravitational acceleration g are length per time squared or LT^{-2}, where T represents the dimension of time. A speed has dimensions of LT^{-1}, so v is a function of g and h with dimensions of LT^{-1}.

Problem 1.4 Dimensions of familiar quantities
In terms of the basic dimensions length L, mass M, and time T, what are the dimensions of energy, power, and torque?

▶ *What combination of g and h has dimensions of speed?*

The combination $\sqrt{gh}$ has dimensions of speed.

$$\Big(\underbrace{\mathrm{LT}^{-2}}_{g}\times\underbrace{\mathrm{L}}_{h}\Big)^{1/2}=\sqrt{\mathrm{L}^2\mathrm{T}^{-2}}=\underbrace{\mathrm{LT}^{-1}}_{\text{speed}}. \tag{1.3}$$

▶ *Is $\sqrt{gh}$ the only combination of g and h with dimensions of speed?*

In order to decide whether $\sqrt{gh}$ is the only possibility, use constraint propagation [43]. The strongest constraint is that the combination of g and h, being a speed, should have dimensions of inverse time (T^{-1}). Because h contains no dimensions of time, it cannot help construct T^{-1}. Because

g contains T^{-2}, the T^{-1} must come from $\sqrt{g}$. The second constraint is that the combination contain L^1. The $\sqrt{g}$ already contributes $L^{1/2}$, so the missing $L^{1/2}$ must come from $\sqrt{h}$. The two constraints thereby determine uniquely how g and h appear in the impact speed v.

The exact expression for v is, however, not unique. It could be $\sqrt{gh}$, $\sqrt{2gh}$, or, in general, $\sqrt{gh} \times$ dimensionless constant. The idiom of multiplication by a dimensionless constant occurs frequently and deserves a compact notation akin to the equals sign:

$$v \sim \sqrt{gh}. \tag{1.4}$$

Including this $\sim$ notation, we have several species of equality:

$$\begin{aligned} &\propto \quad \text{equality except perhaps for a factor with dimensions,} \\ &\sim \quad \text{equality except perhaps for a factor without dimensions,} \\ &\approx \quad \text{equality except perhaps for a factor close to 1.} \end{aligned} \tag{1.5}$$

The exact impact speed is $\sqrt{2gh}$, so the dimensions result $\sqrt{gh}$ contains the entire functional dependence! It lacks only the dimensionless factor $\sqrt{2}$, and these factors are often unimportant. In this example, the height might vary from a few centimeters (a flea hopping) to a few meters (a cat jumping from a ledge). The factor-of-100 variation in height contributes a factor-of-10 variation in impact speed. Similarly, the gravitational acceleration might vary from $0.27\,\mathrm{m\,s^{-2}}$ (on the asteroid Ceres) to $25\,\mathrm{m\,s^{-2}}$ (on Jupiter). The factor-of-100 variation in g contributes another factor-of-10 variation in impact speed. Much variation in the impact speed, therefore, comes not from the dimensionless factor $\sqrt{2}$ but rather from the symbolic factors—which are computed exactly by dimensional analysis.

Furthermore, not calculating the exact answer can be an advantage. Exact answers have all factors and terms, permitting less important information, such as the dimensionless factor $\sqrt{2}$, to obscure important information such as $\sqrt{gh}$. As William James advised, "The art of being wise is the art of knowing what to overlook" [19, Chapter 22].

Problem 1.5 Vertical throw

You throw a ball directly upward with speed v_0. Use dimensional analysis to estimate how long the ball takes to return to your hand (neglecting air resistance). Then find the exact time by solving the free-fall differential equation. What dimensionless factor was missing from the dimensional-analysis result?

1.3 Guessing integrals

The analysis of free fall (Section 1.2) shows the value of not separating dimensioned quantities from their units. However, what if the quantities are dimensionless, such as the 5 and x in the following Gaussian integral:

$$\int_{-\infty}^{\infty} e^{-5x^2}\, dx\,? \tag{1.6}$$

Alternatively, the dimensions might be unspecified—a common case in mathematics because it is a universal language. For example, probability theory uses the Gaussian integral

$$\int_{x_1}^{x_2} e^{-x^2/2\sigma^2}\, dx, \tag{1.7}$$

where x could be height, detector error, or much else. Thermal physics uses the similar integral

$$\int e^{-\frac{1}{2}mv^2/kT}\, dv, \tag{1.8}$$

where v is a molecular speed. Mathematics, as the common language, studies their common form $\int e^{-\alpha x^2}$ without specifying the dimensions of α and x. The lack of specificity gives mathematics its power of abstraction, but it makes using dimensional analysis difficult.

▶ *How can dimensional analysis be applied without losing the benefits of mathematical abstraction?*

The answer is to find the quantities with unspecified dimensions and then to assign them a consistent set of dimensions. To illustrate the approach, let's apply it to the general definite Gaussian integral

$$\int_{-\infty}^{\infty} e^{-\alpha x^2}\, dx. \tag{1.9}$$

Unlike its specific cousin with $\alpha = 5$, which is the integral $\int_{-\infty}^{\infty} e^{-5x^2}\, dx$, the general form does not specify the dimensions of x or α—and that openness provides the freedom needed to use the method of dimensional analysis.

The method requires that any equation be dimensionally valid. Thus, in the following equation, the left and right sides must have identical dimensions:

$$\int_{-\infty}^{\infty} e^{-\alpha x^2}\,dx = \text{something}. \tag{1.10}$$

Is the right side a function of x? Is it a function of α? Does it contain a constant of integration?

The left side contains no symbolic quantities other than x and α. But x is the integration variable and the integral is over a definite range, so x disappears upon integration (and no constant of integration appears). Therefore, the right side—the "something"—is a function only of α. In symbols,

$$\int_{-\infty}^{\infty} e^{-\alpha x^2}\,dx = f(\alpha). \tag{1.11}$$

The function f might include dimensionless numbers such as 2/3 or $\sqrt{\pi}$, but α is its only input with dimensions.

For the equation to be dimensionally valid, the integral must have the same dimensions as $f(\alpha)$, and the dimensions of $f(\alpha)$ depend on the dimensions of α. Accordingly, the dimensional-analysis procedure has the following three steps:

Step 1. Assign dimensions to α (Section 1.3.1).

Step 2. Find the dimensions of the integral (Section 1.3.2).

Step 3. Make an $f(\alpha)$ with those dimensions (Section 1.3.3).

1.3.1 Assigning dimensions to α

The parameter α appears in an exponent. An exponent specifies how many times to multiply a quantity by itself. For example, here is 2^n:

$$2^n = \underbrace{2 \times 2 \times \cdots \times 2}_{n \text{ terms}}. \tag{1.12}$$

The notion of "how many times" is a pure number, so an exponent is dimensionless.

Hence the exponent $-\alpha x^2$ in the Gaussian integral is dimensionless. For convenience, denote the dimensions of α by $[\alpha]$ and of x by $[x]$. Then

$$[\alpha]\,[x]^2 = 1, \tag{1.13}$$

or

$$[\alpha] = [x]^{-2}. \tag{1.14}$$

This conclusion is useful, but continuing to use unspecified but general dimensions requires lots of notation, and the notation risks burying the reasoning.

The simplest alternative is to make x dimensionless. That choice makes α and $f(\alpha)$ dimensionless, so any candidate for $f(\alpha)$ would be dimensionally valid, making dimensional analysis again useless. The simplest effective alternative is to give x simple dimensions—for example, length. (This choice is natural if you imagine the x axis lying on the floor.) Then $[\alpha] = L^{-2}$.

1.3.2 Dimensions of the integral

The assignments $[x] = L$ and $[\alpha] = L^{-2}$ determine the dimensions of the Gaussian integral. Here is the integral again:

$$\int_{-\infty}^{\infty} e^{-\alpha x^2}\, dx. \tag{1.15}$$

The dimensions of an integral depend on the dimensions of its three pieces: the integral sign $\int$, the integrand $e^{-\alpha x^2}$, and the differential dx.

The integral sign originated as an elongated S for *Summe*, the German word for sum. In a valid sum, all terms have identical dimensions: The fundamental principle of dimensions requires that apples be added only to apples. For the same reason, the entire sum has the same dimensions as any term. Thus, the summation sign—and therefore the integration sign—do not affect dimensions: The integral sign is dimensionless.

Problem 1.6 Integrating velocity
Position is the integral of velocity. However, position and velocity have different dimensions. How is this difference consistent with the conclusion that the integration sign is dimensionless?

Because the integration sign is dimensionless, the dimensions of the integral are the dimensions of the exponential factor $e^{-\alpha x^2}$ multiplied by the dimensions of dx. The exponential, despite its fierce exponent $-\alpha x^2$, is merely several copies of e multiplied together. Because e is dimensionless, so is $e^{-\alpha x^2}$.

▶ *What are the dimensions of* dx*?*

To find the dimensions of dx, follow the advice of Silvanus Thompson [45, p. 1]: Read d as "a little bit of." Then dx is "a little bit of x." A little length is still a length, so dx is a length. In general, dx has the same dimensions as x. Equivalently, d—the inverse of $\int$—is dimensionless.

Assembling the pieces, the whole integral has dimensions of length:

$$\left[\int e^{-\alpha x^2}\,dx\right] = \underbrace{\left[e^{-\alpha x^2}\right]}_{1} \times \underbrace{[dx]}_{L} = L. \tag{1.16}$$

> **Problem 1.7 Don't integrals compute areas?**
> A common belief is that integration computes areas. Areas have dimensions of L^2. How then can the Gaussian integral have dimensions of L?

1.3.3 Making an $f(\alpha)$ with correct dimensions

The third and final step in this dimensional analysis is to construct an $f(\alpha)$ with the same dimensions as the integral. Because the dimensions of α are L^{-2}, the only way to turn α into a length is to form $\alpha^{-1/2}$. Therefore,

$$f(\alpha) \sim \alpha^{-1/2}. \tag{1.17}$$

This useful result, which lacks only a dimensionless factor, was obtained without any integration.

To determine the dimensionless constant, set $\alpha = 1$ and evaluate

$$f(1) = \int_{-\infty}^{\infty} e^{-x^2}\,dx. \tag{1.18}$$

This classic integral will be approximated in Section 2.1 and guessed to be $\sqrt{\pi}$. The two results $f(1) = \sqrt{\pi}$ and $f(\alpha) \sim \alpha^{-1/2}$ require that $f(\alpha) = \sqrt{\pi/\alpha}$, which yields

$$\int_{-\infty}^{\infty} e^{-\alpha x^2}\,dx = \sqrt{\frac{\pi}{\alpha}}. \tag{1.19}$$

We often memorize the dimensionless constant but forget the power of α. Do not do that. The α factor is usually much more important than the dimensionless constant. Conveniently, the α factor is what dimensional analysis can compute.

Problem 1.8 Change of variable

Rewind back to page 8 and pretend that you do not know $f(\alpha)$. Without doing dimensional analysis, show that $f(\alpha) \sim \alpha^{-1/2}$.

Problem 1.9 Easy case $\alpha = 1$

Setting $\alpha = 1$, which is an example of easy-cases reasoning (Chapter 2), violates the assumption that x is a length and α has dimensions of L^{-2}. Why is it okay to set $\alpha = 1$?

Problem 1.10 Integrating a difficult exponential

Use dimensional analysis to investigate $\int_0^\infty e^{-\alpha t^3}\,dt$.

1.4 Summary and further problems

Do not add apples to oranges: Every term in an equation or sum must have identical dimensions! This restriction is a powerful tool. It helps us to evaluate integrals without integrating and to predict the solutions of differential equations. Here are further problems to practice this tool.

Problem 1.11 Integrals using dimensions

Use dimensional analysis to find $\int_0^\infty e^{-ax}\,dx$ and $\int \frac{dx}{x^2 + a^2}$. A useful result is

$$\int \frac{dx}{x^2 + 1} = \arctan x + C. \tag{1.20}$$

Problem 1.12 Stefan–Boltzmann law

Blackbody radiation is an electromagnetic phenomenon, so the radiation intensity depends on the speed of light c. It is also a thermal phenomenon, so it depends on the thermal energy $k_B T$, where T is the object's temperature and k_B is Boltzmann's constant. And it is a quantum phenomenon, so it depends on Planck's constant $\hbar$. Thus the blackbody-radiation intensity I depends on c, $k_B T$, and $\hbar$. Use dimensional analysis to show that $I \propto T^4$ and to find the constant of proportionality σ. Then look up the missing dimensionless constant. (These results are used in Section 5.3.3.)

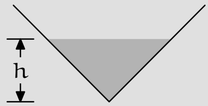
Problem 1.13 Arcsine integral

Use dimensional analysis to find $\int \sqrt{1 - 3x^2}\,dx$. A useful result is

$$\int \sqrt{1 - x^2}\,dx = \frac{\arcsin x}{2} + \frac{x\sqrt{1 - x^2}}{2} + C, \tag{1.21}$$

Problem 1.14 Related rates

Water is poured into a large inverted cone (with a 90° opening angle) at a rate $dV/dt = 10\,\mathrm{m}^3\,\mathrm{s}^{-1}$. When the water depth is $h = 5\,\mathrm{m}$, estimate the rate at which the depth is increasing. Then use calculus to find the exact rate.



Problem 1.15 Kepler's third law

Newton's law of universal gravitation—the famous inverse-square law—says that the gravitational force between two masses is

$$F = -\frac{Gm_1 m_2}{r^2}, \tag{1.22}$$

where G is Newton's constant, m_1 and m_2 are the two masses, and r is their separation. For a planet orbiting the sun, universal gravitation together with Newton's second law gives

$$m\frac{d^2\mathbf{r}}{dt^2} = -\frac{GMm}{r^2}\hat{\mathbf{r}}, \tag{1.23}$$

where M is the mass of the sun, m the mass of the planet, $\mathbf{r}$ is the vector from the sun to the planet, and $\hat{\mathbf{r}}$ is the unit vector in the $\mathbf{r}$ direction.

How does the orbital period τ depend on orbital radius r? Look up Kepler's third law and compare your result to it.