Completing the Square: Beyond the Quadratic Formula

Jeffrey M. Rabin
jrabin@ucsd.edu

Dept. of Mathematics, UCSD

Curtis Center Mathematics and Teaching Conference
UCLA, March 5, 2016
Since algebra surpasses all human subtlety and the clarity of every mortal mind, it must be accounted a truly celestial gift, which gives such an illuminating experience of the true power of the intellect that whoever attains to it will believe there is nothing he cannot understand.

-Ars Magna, Girolamo Cardano (1501-1576).
Completing the square in the derivation of the quadratic formula
Completing the cube to (try to) solve the cubic equation
Quadratic maximum/minimum problems
The arithmetic-geometric mean inequality, and more max/min problems
Tangent lines without calculus
Quadratic forms in several variables
Must a positive polynomial be a sum of squares?
Completing the square is not an isolated trick but a way of thinking about quadratic functions in many contexts.
Al-kitab al-mukhtasar fi hisab al-jabr wal-muqabala
(The Compendious Book on Calculation by Completion and Balancing),
Muhammad ibn Musa Al-Khwarizmi, circa 825 A.D.
A method for *transforming* a quadratic function 
\[ f(x) = ax^2 + bx + c \] to make its zeros and symmetry evident.

Compare the first two terms of 
\[ f(x) = a\left(x^2 + \frac{b}{a}x + \frac{c}{a}\right) \]
with those of the perfect square 
\[ (x + q)^2 = x^2 + 2qx + q^2. \]

\[ f(x) = a\left[(x + \frac{b}{2a})^2 + \frac{c}{a} - \frac{b^2}{4a^2}\right] = a\left[(x + \frac{b}{2a})^2 - \frac{b^2-4ac}{4a^2}\right]. \]

If our aim is simply to solve the equation 
\[ f(x) = 0 \]
we can divide by \( a \) and get 
\[ (x + \frac{b}{2a})^2 = \frac{b^2-4ac}{4a^2}. \]

It is easier to find \( y = x + \frac{b}{2a} \) than to find \( x \) directly.
Any quadratic equation can be transformed into the simple form $x^2 = \text{some number}$.

The function $x^2$ has two key properties (for real $x$): symmetry $(-x)^2 = x^2$ and positivity $x^2 \geq 0$.

These account for the symmetry axis of a parabola and the fact that its vertex is a minimum or maximum point.

No real solution to $f(x) = 0$ if $b^2 - 4ac < 0$, because the minimum (maximum) lies above (below) the $x$ axis.

Algebra is transformation and deduction. (Or is it completion and balancing?)
The Classic Quadratic Equation Problem

- Find two numbers given their sum and product. (Babylonian, 1700 B.C.?)
- If \( x + y = S \) and \( xy = P \) we get \( x(S - x) = P \) or \( x^2 - Sx + P = 0 \).
- This is familiar from the instructions we give our students for factoring such a quadratic!

Solution \( x = (S \pm \sqrt{S^2 - 4P})/2 \), or, in terms of the average \( A = S/2 \) of the two numbers, \( x = A \pm \sqrt{A^2 - P} \). The two signs actually give both numbers!

- The two numbers are real if \( P \leq A^2 \) and complex otherwise.
- This is a disguised version of the arithmetic-geometric mean inequality, more conventionally written \( \sqrt{P} \leq A \): the geometric mean of two (positive) real numbers is never greater than the arithmetic mean, and is equal to it only when the two numbers are equal.

JMR

Completing the Square
Attempt the same strategy for solving $ax^3 + bx^2 + cx + d = 0$.

May as well divide both sides by $a$ immediately. Equivalently, assume $a = 1$ and solve $x^3 + bx^2 + cx + d = 0$.

Compare the first two terms with those of the perfect cube $(x + q)^3 = x^3 + 3qx^2 + 3q^2x + q^3$. They match if $q = b/3$.

Equation becomes $(x + q)^3 - 3q^2x - q^3 + cx + d = 0$.

The $x^2$ term has been removed but the $x$ term generally remains.

In terms of the new variable $y = x + q$ we still have an equation of the form $y^3 + ry + s = 0$ to solve.
In $y^3 + ry + s = 0$, let $y = u + v$.

$(u + v)^3 + r(u + v) + s = 0$ simplifies to $u^3 + v^3 + (3uv + r)(u + v) + s = 0$.

Wouldn’t it be great if $3uv + r = 0$? That is, $uv = -r/3$.

Then the problem is to find $u, v$ given that $u^3v^3 = -r^3/27$ and $u^3 + v^3 = -s$.

We can find $u^3, v^3$ from the classic quadratic problem of finding two numbers given their sum and product!

Then take the cube root to get $u$ and $v$, hence $y = u + v$.

But wait! This gives a single solution, and a cubic equation should (generally) have three!

The number $u^3$ has three complex cube roots, and we need all of them to get all three solutions! (Even if all are real!)

This was historically a huge motivation for taking complex numbers seriously.
Quadratic Max/Min Problems

- More interesting for students (?) than solving quadratic equations.
- Realistic applications and unexpected solutions.
- Promote thinking in terms of functions rather than single unknowns.
- Archetypal example: What rectangle with given perimeter has the greatest area?
  - If the sides are $x, y$ then giving the perimeter is equivalent to giving the sum $x + y = S$.
  - The area is $A(x) = x(S - x) = Sx - x^2 = (S^2/4) - (x - \frac{S}{2})^2$.
  - Because perfect squares are nonnegative, the maximum $A$ is $S^2/4$ and occurs when the sides are equal: $x = y = S/2$.
  - Area increases steadily as the rectangle becomes more square.
The Arithmetic-Geometric Mean Inequality

- We have proved (again!) the AGM inequality: if the sum of two positive numbers is given, then their maximum possible product occurs when they are equal, and is the square of their average.

\[ \sqrt{xy} \leq \frac{(x + y)}{2}. \]

- A disguised version of this principle: which product is larger, \((45)(87)\) or \((47)(85)\)?

JMR

Completing the Square
The Arithmetic-Geometric Mean Inequality

- We have proved (again!) the AGM inequality: if the sum of two positive numbers is given, then their maximum possible product occurs when they are equal, and is the square of their average.
  \[ \sqrt{xy} \leq \frac{x + y}{2}. \]

- A disguised version of this principle: which product is larger, (45)(87) or (47)(85)?

- The latter, since that pair is closer to their average 66.

- The AGM inequality can also be read this way: if the product of two positive numbers is given, then the minimum possible sum occurs when they are equal.

- Quadratic max/min problems can be solved directly by completing the square, or by using this inequality.
Another Max/Min Problem

- A wire of given length (say 1 meter) is cut into two parts. One part is bent into the shape of a circle, the other into a square. What are the maximum and minimum possible combined areas of the circle and square?
Another Max/Min Problem

- A wire of given length (say 1 meter) is cut into two parts. One part is bent into the shape of a circle, the other into a square. What are the maximum and minimum possible combined areas of the circle and square?
- If the segments have lengths $x$ and $1 - x$ then the circle has radius $x/2\pi$ and the square has side $(1 - x)/4$. Then the sum of their areas is
  \[ A(x) = \frac{x^2}{4\pi} + \frac{(1-x)^2}{16}, \]
  or
  \[ 16A(x) = \frac{\pi+4}{\pi} x^2 - 2x + 1. \]
- The graph is a parabola opening upward with vertex at
  \[ x = -\frac{b}{2a} = \frac{\pi}{\pi+4} \approx 0.44, \]
  giving the minimum area $A \approx 0.035$.
- For $0 \leq x \leq 1$ the maximum must be at an endpoint $x = 0$ or $x = 1$, assuming it is legal to “cut” one piece of wire having zero length.
- $A(0) = 0.0625$ (all square); $A(1) \approx 0.07958$ (all circle).
Total Area in the Wire Cutting Problem

In[3]:= graph \( y = \frac{2.27x^2 - 2x + 1}{16} \) for \( 0 < x < 1 \) and \( 0 < y < 0.1 \)

Assuming “0” is referring to math | Use as a decimal number instead

Input interpretation:

| plot | \( \frac{1}{16} (2.27x^2 - 2x + 1) \) | \( x = 0 \) to \( 1 \)
| \( y = 0 \) to \( 0.1 \) |

Plot:

Arc length of curve:
\[ 1.01563 - 0.0709375 x + 0.0805141 x^2 \quad \text{at} \quad x = 1.00349 \]
A jolly rancher has a given length $F$ of fencing available to enclose a rectangular corral. He also wants to subdivide it into four separate enclosures with more fencing parallel to one of the sides. What is the maximum area he can enclose?

$$2l + 5w = F$$
Call the length $l$ and the width $w$; suppose the additional fencing is parallel to the width. Then the total length of fencing is $2l + 5w$ and this is given.

The area is $lw = (1/10)(2l)(5w)$. So we are maximizing the product of numbers whose sum is given!

The maximum occurs when $2l = 5w$, each being half the given length of fencing.

The maximum area is $(1/10)(F^2/4)$, as compared to the area $(1/4)(F^2/4)$ that could be achieved if the internal subdivision wasn’t necessary, that is, $2/5$ as much.

The AGM approach explains why the maximum has $2l = 5w$, which tends to look coincidental in the approach by completing the square, or calculus. It is also clearer how the answer changes if more separate enclosures are needed.
Let $a_1 \leq a_2 \leq a_3 \leq \cdots \leq a_n$ be any list of positive real numbers, arranged for convenience in increasing order. Their \textit{arithmetic mean} is $A = (a_1 + a_2 + \cdots + a_n)/n$, and their \textit{geometric mean} is $G = (a_1 a_2 \cdots a_n)^{1/n}$.

Then, remarkably, once again $G \leq A$ with equality only when all the numbers are equal.

This again tells us the maximum product of a list of numbers whose sum is given, or the minimum sum when the product is given!
What is the minimum value of $6x + \frac{24}{x^2}$ for positive $x$?

Rewriting the expression as $3x + 3x + \frac{24}{x^2}$, it is now the sum of three positive numbers whose product is $3 \times 3 \times 24 = 216 = 6^3$, so $G = 6$.

This is minimized when $3x = 3x = \frac{24}{x^2} = 6$, so the minimum sum is 18, when $x = 2$.

Very powerful, and explains the result rather than making it look like a coincidence.

Max/min problems in calculus often have a solution where different terms “balance”, for no evident reason.
Suppose $A$ is given, how large can we make $G$?

If the numbers $a_1 \leq a_2 \leq \cdots \leq a_n$ are not all equal already, then $a_1 < A < a_n$. If we increase $a_1$ to the new value $A$, and simultaneously decrease $a_n$ by the same amount $A - a_1$, we do not alter their sum, but we increase their product by moving them closer together. That is, we have increased $G$ while keeping $A$ fixed.

We can continue to do this as long as the numbers are not all equal to $A$. (Re-order them into increasing order after each step.) Thus $G$ is maximized when all the numbers are equal to $A$, and at that point clearly $G = A$ too.

Thus, as claimed, $G \leq A$ with equality only when the numbers are all equal.
Another Example Normally Done By Calculus

- A cylindrical aluminum can of radius $r$ and height $h$ has volume $V = \pi r^2 h$ and surface area $S = 2\pi r^2 + 2\pi rh$.
- The manufacturer would naturally like to minimize $S$ for a given value of $V$. What dimensions will achieve this?
A cylindrical aluminum can of radius $r$ and height $h$ has volume $V = \pi r^2 h$ and surface area $S = 2\pi r^2 + 2\pi rh$.

The manufacturer would naturally like to minimize $S$ for a given value of $V$. What dimensions will achieve this?

Write the surface area as $2\pi r^2 + \pi rh + \pi rh$, the sum of three numbers whose product is $2\pi^3 r^4 h^2 = 2\pi V^2$. Then the minimum occurs when $2\pi r^2 = \pi rh$, in other words $2r = h$.

Minimum $S = 3(2\pi V^2)^{1/3}$.

Check out the cans the next time you’re in the market.
A Problem for You

You are invited to choose a slice of pie (in the usual shape, a circular sector) with any central angle you like. However, the radius of your slice will be adjusted so that its perimeter is a predetermined number, maybe 8 inches. What angle do you choose in order to get the maximum-area slice?
A line \( y = mx + k \) intersects a parabola \( y = ax^2 + bx + c \) at points where \( ax^2 + bx + c = mx + k \), or 
\[
ax^2 + (b - m)x + (c - k) = 0.
\]
There may be two, one or zero solutions to this equation. If there are two or none, the line is not tangent to the parabola: we want just one solution.

That is, we want this quadratic to be a perfect square! (Or, zero discriminant.)

The only perfect square that matches the first two terms is of course

\[
a \left( x + \frac{b - m}{2a} \right)^2.
\]

so that must be it!

The tangency occurs at \( x = (m - b)/2a \), in other words \( m = 2ax + b \).

This is exactly the calculus formula for the slope of a tangent line at point \( x \).
Setting the discriminant to zero is not as informative:

\[(b - m)^2 - 4a(c - k) = 0\]

is the condition for the line \(y = mx + k\) to be tangent to the parabola somewhere, but we don’t know where. For example, the lines tangent to \(y = x^2\) somewhere all have \(m^2 + 4k = 0\).
The previous approach generalizes to the graph of any polynomial function $y = f(x)$. A line $y = mx + b$ intersects this graph when $f(x) - mx - b = 0$. If $x = x_0$ is a solution of this equation, $b = f(x_0) - mx_0$ and the left side is divisible by $x - x_0$.

The above holds for any intersection point. A point of tangency is special because the left side has two factors of $x - x_0$, that is, it has the perfect square factor $(x - x_0)^2$. 
For example, let’s find the tangent line to $y = x^3 - 3x + 2$ at the point $(2, 4)$. According to the above reasoning, the equation $x^3 - 3x + 2 - mx - b = 0$ should have two factors of $x - 2$. Synthetic division is an efficient tool for checking this.

Dividing once by $x - 2$ we find the remainder $4 - b - 2m$. Assuming this to be zero and dividing again produces the remainder $9 - m$ and the quotient $x + 4$. Setting the remainders equal to zero tells us the slope $m = 9$ and intercept $b = 4 - 2m = -14$ of the tangent line.

The cubic factors into $(x - 2)^2(x + 4)$. The last factor tells us there is another intersection at $x = -4$.

This works on any polynomial, all without calculus!
Tangent Line $y = 9x - 14$ to Cubic $y = x^3 - 3x + 2$
Let \( Q(x, y, z) = x^2 + 5y^2 + z^2 - 2xy + 2yz + 2xz \), a quadratic form in three variables. Evidently \( Q(0, 0, 0) = 0 \). If \( Q(x, y, z) > 0 \) for all other values of \( x, y, z \) then \( Q \) has its minimum value at \( (0, 0, 0) \). How can we determine whether this is true?

Random guessing of points, e.g. \( Q(1, 1, 1) = 9 \), may or may not reveal any negative values. A systematic method is to complete the square, in alphabetical order! Pretend for a moment that only \( x \) is a variable, while \( y \) and \( z \) are constants.
Multivariable Quadratic Forms

- \( Q(x, y, z) = x^2 + 5y^2 + z^2 - 2xy + 2yz + 2xz \)
- \( Q = x^2 + 2(z - y)x + 5y^2 + z^2 + 2yz = (x + z - y)^2 + 4y^2 + 4yz \). Next, in the last two terms, pretend that only \( y \) is a variable, and get
  - \( Q = (x - y + z)^2 + (2y + z)^2 - z^2 \). Now it is clear that \( Q \) is sometimes negative, and we can see how to find an example. \( Q(6, 2, -4) = -16 \). So (0, 0, 0) is not a minimum point.

- If one completes the square in reverse alphabetical order \( z, y, x \), one does not get the same answer, but rather \( Q = (x + y + z)^2 + (2y - x)^2 - x^2 \). However, the signature of \( Q \), the fact that two squares are positive and one negative, is always the same!
Hilbert’s Seventeenth Problem

- Evidently one way to show that a function is nonnegative is to write it as the sum of squares of other expressions. Is this always possible?
- Specifically, if a polynomial in several variables has all its values nonnegative, does it follow that it can be written as the sum of squares of other polynomials?
- The answer is NO; let \( P(x, y) = x^4y^2 + x^2y^4 + 1 - 3x^2y^2 \). Then \( P(x, y) \geq 0 \) for all values of \( x, y \): an equivalent way to write this claim is

\[
\left(\frac{1}{3}\right)(x^4y^2 + x^2y^4 + 1) \geq x^2y^2,
\]

and this is an example of the AGM inequality. However, there is no way to write \( P(x, y) \) as the sum of several polynomials squared.

- Such squares would have the form \( (a + bxy + cx^2y + dxy^2)^2 \), but this gives \( x^2y^2 \) with a positive coefficient.
Hilbert’s Seventeenth Problem

Hilbert knew this, and his 17th problem (of 23, in 1900) asked instead whether every nonnegative polynomial could be written as the sum of squares of *rational functions*. The answer is YES, as proved by Emil Artin in 1927. In the previous example, one can check that \( P(x, y) = x^4y^2 + x^2y^4 + 1 - 3x^2y^2 \) is the sum of squares

\[
\frac{x^2y(x^2 + y^2 - 2)}{x^2 + y^2} + \frac{xy^2(x^2 + y^2 - 2)}{x^2 + y^2} + \frac{xy(x^2 + y^2 - 2)}{x^2 + y^2} + \frac{(x^2 - y^2)}{x^2 + y^2}
\]

There are algorithms, in the same spirit as completing the square but more complex, that will determine whether a polynomial is indeed nonnegative, and, if so, find these rational functions (or polynomials, when that is possible).

Important for computer solution of complex max/min problems.
If a polynomial has only positive values, must it have a minimum value?

Certainly YES for polynomials \( f(x) \) in ONE variable. Those must have even degree, and be very large outside some interval, with a finite number of local minima where \( f'(x) = 0 \).

\[ \int_{-6}^{5} \sqrt{1 + (40 + 36x - 9x^2 - 4x^3)^2} \, dx \approx 1013.67… \]
But NO in general, for example \( f(x, y) = (xy - 1)^2 + x^2 \).

This is positive, never zero, and can be made as small as desired by choosing \( x \) small and \( y = 1/x \).

0 is not called a *minimum value* unless \( f(x, y) = 0 \) at some specific point. In this example it is the *greatest lower bound* of \( f(x, y) \).
Algebra is transformation and deduction (OK, and completion and balancing).

Perfect squares have the key properties of symmetry \((-x)^2 = x^2\) and positivity \(x^2 \geq 0\). Completing the square is useful to bring out these properties whenever quadratic functions appear.

Inequalities, max/min problems, tangent lines.

Second derivative test is based on completing the square.

Standards for Mathematical Practice: Look for and make use of structure; Look for and express regularity in repeated reasoning.
References

Since algebra surpasses all human subtlety and the clarity of every mortal mind, it must be accounted a truly celestial gift, which gives such an illuminating experience of the true power of the intellect that whoever attains to it will believe there is nothing he cannot understand.

-Ars Magna, Girolamo Cardano (1545).

jrabin@ucsd.edu
The optimal angle is $\theta = 2$ radians, about 115 degrees. The resulting slice of pie has as much area as a square with the same perimeter would have. However, one can vary the angle considerably, say $1 < \theta < 3$ radians, losing no more than a few percent of this area.