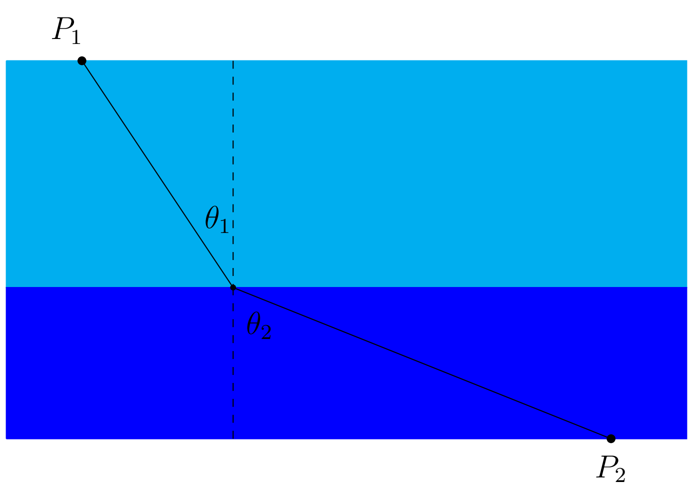
# Math 1131 Applications: Optimization, I

Optimization problems occur in many areas, as shown here. We discuss here a few such problems that take too long to set up and solve in lecture.

**Application 1**. Snell's law.

Light rays bend when traveling from one medium to another, such as air to water. This causes items placed in or behind water to look shifted, as in the images below.

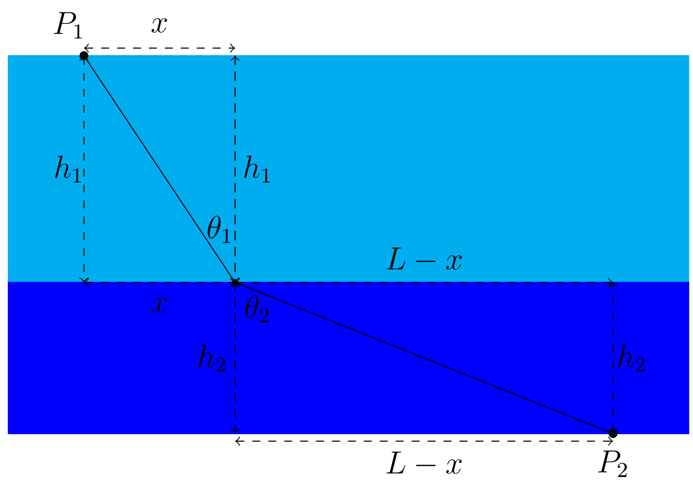
The bending of a light ray from $P_1$ in air to $P_2$ in water (see below) fits Snell's law. It says $\boxed{\dfrac{\sin\theta_1}{\sin\theta_2} = \dfrac{v_1}{v_2}}$, where light's path from $P_1$ to water has angle $\theta_1$, light's path in water to $P_2$ has angle $\theta_2$, light has speed $v_1$ in air and speed $v_2$ in water.




The physical basis for Snell's law is not that light takes the path from $P_1$ to $P_2$ of *least distance*, since that would be a straight line ($\theta_1 = \theta_2$), but instead that light follows a path of *least time*: the two straight lines along which a light ray travels from $P_1$ to $P_2$ are those for which the combined time along the paths is minimal. Let's use calculus to show the path of least time from $P_1$ to $P_2$ satisfies Snell's law.

We need a few more variables besides the angles, as in the figure below. Let

- $x$ be the distance from $P_1$ to the line perpendicular to the boundary that passes through the point where a light ray from $P_1$ meets the boundary,
- $h_1$ be the vertical distance from $P_1$ to the boundary,
- $h_2$ be the vertical distance from $P_2$ to the boundary,
- $L$ be the total horizontal separation between $P_1$ and $P_2$: $L = x + (L - x)$.




Let $v_1$ be the speed of light in air and $v_2$ be the speed of light in water. Let the light ray from $P_1$ to $P_2$ take time $t_1$ to travel from $P_1$ to the boundary and time $t_2$ to travel from the boundary to $P_2$. We can compute the distance from $P_1$ to the point where the light ray from $P_1$ reaches the boundary in two ways: by the distance formula with right triangles and by "rate times time":

$$\sqrt{x^2 + h_1^2} = v_1 t_1.$$

Using $P_2$ in place of $P_1$,

$$\sqrt{(L-x)^2 + h_2^2} = v_2 t_2.$$

Therefore the *total time* for the light ray to travel from $P_1$ to $P_2$ is

$$t_1 + t_2 = \frac{1}{v_1}\sqrt{x^2 + h_1^2} + \frac{1}{v_2}\sqrt{(L-x)^2 + h_2^2}.$$

We want an $x$ in $[0, L]$ that minimizes the total time. The only flexible parameter is $x$, so write the total time above as a function of $x$, say $T(x)$. Its $x$-derivative is

$$T'(x) = \frac{1}{v_1}\frac{2x}{2\sqrt{x^2+h_1^2}} + \frac{1}{v_2}\frac{2(L-x)(-1)}{2\sqrt{(L-x)^2+h_2^2}} = \frac{1}{v_1}\frac{x}{\sqrt{x^2+h_1^2}} - \frac{1}{v_2}\frac{L-x}{\sqrt{(L-x)^2+h_2^2}},$$

which is continuous for $0 \leq x \leq L$. Since $T'(0) = -L/(v_2\sqrt{L^2+h_2^2}) < 0$ and $T'(L) = L/(v_1\sqrt{L^2+h_1^2}) > 0$, we must have $T'(x) = 0$ for some $x$ in $(0, L)$. In fact there is *only one* such $x$ since $T'$ is increasing on account of $T'$ having a positive derivative: check with the quotient rule that for numbers $h > 0$,

$$\frac{d}{dx}\frac{x}{\sqrt{x^2+h^2}} = \frac{h^2}{(x^2+h^2)^{3/2}}, \quad \frac{d}{dx}\frac{L-x}{\sqrt{(L-x)^2+h^2}} = \frac{-h^2}{((L-x)^2+h^2)^{3/2}},$$

so

$$\begin{aligned} T''(x) &= \frac{1}{v_1}\frac{h_1^2}{(x^2+h_1^2)^{3/2}} - \frac{1}{v_2}\frac{-h_2^2}{((L-x)^2+h_2^2)^{3/2}} \\ &= \frac{1}{v_1}\frac{h_1^2}{(x^2+h_1^2)^{3/2}} + \frac{1}{v_2}\frac{h_2^2}{((L-x)^2+h_2^2)^{3/2}} \\ &> 0. \end{aligned}$$

At the critical number $x$ in $(0, L)$ where $T'(x) = 0$ we have

$$\frac{1}{v_1}\frac{x}{\sqrt{x^2+h_1^2}} = \frac{1}{v_2}\frac{L-x}{\sqrt{(L-x)^2+h_2^2}}.$$

In terms of trigonometry, the left side is $(1/v_1)\sin\theta_1$ and the right side is $(1/v_2)\sin\theta_2$, so $(1/v_1)\sin\theta_1 = (1/v_2)\sin\theta_2$, which is the same as $(\sin\theta_1)/(\sin\theta_2) = v_1/v_2$. This path where $T'(x) = 0$ is a *minimum* of $T(x)$, not a maximum, by the second derivative test since $T''(x) > 0$. Thus a light ray from $P_1$ to $P_2$ that moves along a path of *least time* satisfies Snell's law $\dfrac{\sin\theta_1}{\sin\theta_2} = \dfrac{v_1}{v_2}$.

Snell's law (based on paths of least time) explains the appearance of a mirage on a hot road (with $\theta_2 = \pi/2$), behavior of sound waves across a lake and shock waves moving through different layers of rock.

**Application 2**. The Kelly criterion (a long-term fractional betting strategy)

Suppose a coin has a probability $p$ to come up heads and a probability $q = 1 - p$ to come up tails, so a fair coin has $p = q = 1/2$ and a biased coin has $p$ and $q$ not equal to $1/2$. You will bet on the coin coming up heads ("win") by using a *fractional* *betting strategy*, which means betting a fixed fraction $f$ ($0 \le f \le 1$) of your gambling money on heads each time. The amount of money available for gambling is called your *bankroll*, even if you don't use all of it on each bet.

**Example**. Three coin flips come up HHT. Starting with \$100 and using $f = 1/4$, your bankroll after each coin flip is \$100 $\to$ \$125 $\to$ \$156.25 $\to$ \$117.19 (after rounding). If instead $f = 1/2$ (bet half of what you have on heads each time) then your bankroll after each flip is \$100 $\to$ \$150 $\to$ \$225 $\to$ \$112.50. If instead $f = 1$ (bet everything you have on heads each time), then you lose all your money the first time the coin comes up tails, which is a bad long-term strategy unless the coin always comes up heads.

The optimization question here is: in terms of the coin flip probabilities $p$ (for heads), what fractional amount $f$ will maximize the long-term expected earnings with this betting strategy?

Winning a bet increases a bankroll $\$B$ to $\$B + f\$B = \$((1+f)B)$ and losing a bet decreases a bankroll $\$B$ to $\$B - f\$B = \$((1-f)B)$. Every time you bet the fraction $f$ of your bankroll, it changes by the factor $1 + f$ when you win and by the factor $1 - f$ when you lose. Therefore when you bet the fraction $f$ of your bankroll on heads for $n$ successive coin flips starting with $\$B_0$, your bankroll $\$B_n$ after $n$ coin flips is given by the formula

$$\$B_n = \$(1+f)^{W_n}(1-f)^{L_n}B_0,$$

where the coin is heads ("win") $W_n$ times and tails ("lose") $L_n$ times: $W_n + L_n = n$. Since the probability of the coin flip being heads is $p$ and being tails is $q$, for large $n$ we have $W_n/n \approx p$ and $L_n/n \approx q$, so $W_n \approx pn$ and $L_n \approx qn$. Using these in the exponents in the formula for $B_n$ above,

$$B_n \approx (1+f)^{pn}(1-f)^{qn}B_0 = ((1+f)^p(1-f)^q)^n B_0.$$

This suggests that to maximize the bankroll $\$B_n$ after $n$ coin flips for large $n$, we choose $f$ in $[0,1]$ to maximize $(1+f)^p(1-f)^q$. We finally have a calculus problem:

Find the number $f$ in $[0,1]$ that maximizes $G(f) = (1+f)^p(1-f)^q$.

The factor $G(f)$ is called the "gain," which explains the notation $G$ we use here. That the variable of this function is written as $f$ might feel peculiar (isn't $f$ always a function?), but it is a commonly used notation in this problem and it's important to get comfortable working with functions where the variable is not always written as $x$.

To apply the first derivative test to maximize $G(f)$, we compute $G'(f)$:

$$\begin{aligned} G'(f) &= p(1+f)^{p-1}(1-f)^q - (1+f)^p q(1-f)^{q-1} \\ &= (1+f)^{p-1}(1-f)^{q-1}(p(1-f) - q(1+f)) \\ &= (1+f)^{p-1}(1-f)^{q-1}(p-q-(p+q)f). \end{aligned}$$

The factors $(1+f)^{p-1}$ and $(1-f)^{q-1}$ are both positive[1] so $G'(f) = 0$ only when $p - q - (p+q)f = 0$. Since $p + q = 1$,

$$G'(f) = 0 \iff p - q - f = 0 \iff f = p - q = p - (1-p) = 2p - 1.$$

It is left to the reader to check the critical number $f = 2p - 1$ is a maximum, not a minimum, by showing $G''(2p-1) = -p^{p-1}q^{q-1}/2 < 0$.

Using the fractional betting strategy with $f = 2p - 1$ is called the Kelly criterion. It should perform better *in the long run* than any other fractional betting strategy, and of course depends on knowing $p$!

**Example**. If $p = .50$ then the Kelly criterion says $f = 0$: the optimal fractional betting strategy on a fair coin is to bet nothing (the bet has no "edge" to exploit).

**Example**. If $p = .60$ (so $q = .40$) then the Kelly criterion says $f = 2p - 1 = .20$: bet 20% of your bankroll.

**Example**. If $p = .75$ (so $q = .25$) then the Kelly criterion says $f = 2p - 1 = .50$: bet half of your bankroll.

**Example**. If $p < .50$ then the Kelly criterion says $f = 2p - 1 < 0$. What is a negative fractional bet? What $f < 0$ means is that you should use a fractional betting strategy on heads *not* coming up (that is, bet on tails) when heads is less likely than tails. That should make intuitive sense.

**Remark**. The treatment of the Kelly criterion above assumes the amount won or lost on a bet is the amount that was bet: if you bet \$20 and win then you receive \$20, while if you lose the bet you go down by \$20. This doesn't cover the case of bets that come with *odds*, like a 2-to-1 bet where you double your money if you win and

[1] We could have $1 - f = 0$ if $f = 1$, but that fractional betting strategy only makes sense here when the coin always comes up heads, and such a sure thing is not a realistic setting where anyone would be offered a chance to bet, so we will ignore it.

lose what you bet if you lose. If the betting is structured so that the amount you win is $w$ times what you bet and the amount you lose is $\ell$ times what you bet (so 2-to-1 odds means $w = 2$ and $\ell = 1$ and betting without odds means $w = \ell = 1$), then a win on bet of $\$fB$ changes your bankroll to $\$(B + wfB) = \$(1 + wf)B$ and a loss changes your bankroll to $\$(B - \ell fB) = \$(1 - \ell f)B$. The bankroll after $n$ coin flips is

$$B_n \approx (1 + wf)^{pn}(1 - \ell f)^{qn}B_0 = ((1 + wf)^p(1 - \ell f)^q)^n B_0,$$

so you want to find $f$ in $[0, 1]$ that maximizes $G(f) = (1 + wf)^p(1 - \ell f)^q$. It is left to the reader to show this occurs at

$$f = \frac{pw - q\ell}{w\ell} = \frac{p}{\ell} - \frac{q}{w}.$$

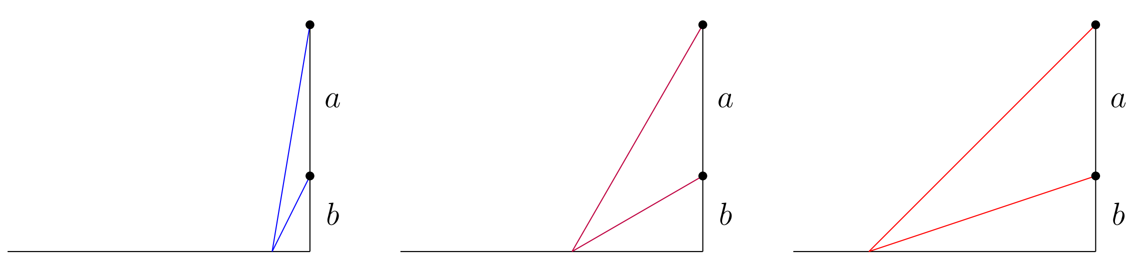
This is the Kelly criterion for betting with odds that are no longer 1-to-1. When $w = \ell = 1$ this formula becomes our earlier formula $f = p - q = 2p - 1$. (We have $f > 0$ when $pw > q\ell$. If $pw < q\ell$ then $f < 0$ and you should bet on tails coming up.)

**Example**. If $p = .60$ (so $q = .40$), $w = 2$, and $\ell = 1$ (2-to-1 odds) then $f = p - q/2 = .60 - .20 = .40$, so bet 40% of your bankroll. In the previous example with $p = .60$ and 1-to-1 odds, the Kelly criterion is to bet 20% of your bankroll. It makes sense that when the odds on heads are 2-to-1, the fractional amount to bet goes up.

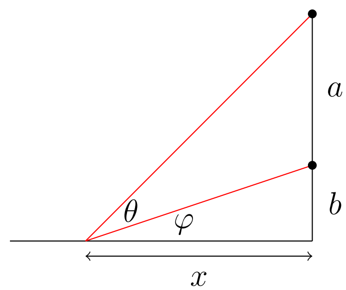
**Application 3**. Viewing angle maximization problem.

Suppose a painting (or movie screen) that is $a$ feet tall is placed on a wall starting $b$ feet from the ground, as shown below. We want the distance from the wall, in terms of $a$ and $b$, that makes the viewing angle from the ground to the painting (or movie screen) maximal. This problem is due to Regiomontanus in the 1400s.

Since the viewer is *below* the lowest level of the screen, the viewing angle very close to the wall is small and the viewing angle very far from the wall is small, so the largest viewing angle is somewhere in between. Where is it?



At a distance $x$ from the wall, let $\theta$ be the viewing angle (see figure below). If $x$ is very small or very large then $\theta$ is very small, so there is some $x$ where $\theta$ is maximal.




Rather than getting a formula for $\theta$ in terms of $x$ (and $a$ and $b$) we will get a formula for $\tan\theta$. At distance $x$ from the wall let $\varphi$ be the angle from the ground to height $b$ on the wall, so by using right triangles, $\tan\varphi = b/x$ and $\tan(\theta+\varphi) = (a+b)/x$. We can get $\tan\theta$ from these using the addition formula $\tan(\alpha+\beta) = \dfrac{\tan\alpha + \tan\beta}{1 - \tan\alpha\tan\beta}$:

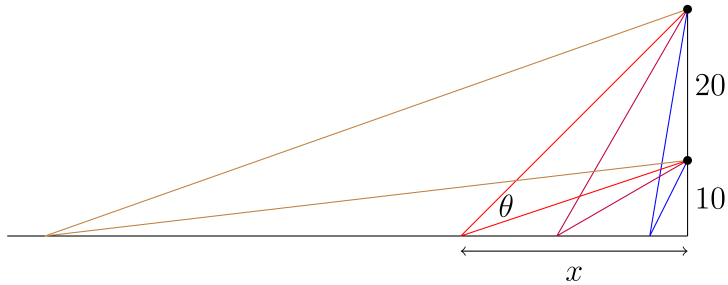
$$\begin{aligned}
\tan\theta &= \tan((\theta+\varphi) - \varphi) \\
&= \frac{\tan(\theta+\varphi) + \tan(-\varphi)}{1 - \tan(\theta+\varphi)\tan(-\varphi)} \quad \text{using } \alpha = \theta+\varphi, \beta = -\varphi \\
&= \frac{\tan(\theta+\varphi) - \tan\varphi}{1 + \tan(\theta+\varphi)\tan\varphi} \quad \text{since } \tan(-\varphi) = -\tan\varphi \\
&= \frac{(a+b)/x - b/x}{1 + ((a+b)/x)(b/x)} \\
&= \frac{ax}{x^2 + (a+b)b}.
\end{aligned}$$

We want to maximize $\theta$ as $x$ varies over positive numbers. Differentiate both sides with respect to $x$:

$$(\sec^2\theta)\frac{d\theta}{dx} = \frac{a(x^2 + (a+b)b) - ax(2x)}{(x^2 + (a+b)b)^2} = \frac{a((a+b)b - x^2)}{(x^2 + (a+b)b)^2}.$$

The term $\sec^2\theta$ is positive. Since $x > 0$ we have $d\theta/dx = 0 \iff x = \sqrt{(a+b)b}$. If $0 < x < \sqrt{(a+b)b}$ then $d\theta/dx > 0$, and if $x > \sqrt{(a+b)b}$ then $d\theta/dx < 0$, so at $\boxed{x = \sqrt{(a+b)b}}$ the angle $\theta$ is maximal.

**Example**. A movie screen 20 feet high is placed on a wall starting at the 10 foot mark. At what distance from the wall, in feet, is the viewing angle towards the screen maximal?




From the work above, the maximal viewing angle occurs when $x = \sqrt{(20+30)10} = \sqrt{500} \approx 22.36$ feet.

**Application 4**. Optimization is not only something done to functions, but also to shapes. A Numberphile video about one such problem, called the "moving sofa problem" (see screenshot below) and still unsolved, is here. During 5:30-6:00 of the video, the speaker describes a shape being "locally optimal" and that is analogous to a local maximum or local minimum of a function. If you watch the whole video you'll find an unexpected connection to the movie Zoolander.