109 Example (Lagrange Multipliers)
Minimize

$$f(x,y) = x^2 + y^2$$

subject to

$$g(x,y) = 2x + y - 10 = 0.$$

The Lagrangian function is

$$L(x,y,\lambda) = x^2 + y^2 + \lambda(2x + y - 10).$$

Setting all partial derivatives to zero implies the maximum must satisfy the equations

$$\begin{aligned} \frac{\partial L}{\partial x} &= 2x + 2\lambda = 0 \\ \frac{\partial L}{\partial y} &= 2y + \lambda = 0 \\ \frac{\partial L}{\partial \lambda} &= 2x + y + 10 = 0 \end{aligned}$$

Solving the first two equations for x and y and substituting into the last equation and solving for λ implies $x = 4$, $y = 2$ and $\lambda = 4$.

Measuring Information

What is information and how do we measure it?

110 Example (Coin Toss)
Assume each of the six coins shown below is displaying "heads."[14]

[14] Image URL:
https://izbicki.me/blog/how-to-create-an-unfair-coin-and-prove-it-with-math.html

(a) I toss one of the six coins and inform you that the coin toss outcome is heads. For which coin will I have given you the most information.
(b) If I inform you the outcome is tails, for which coin would I have given you the most information.
(c) If I toss a two-headed coin *and you know that it is two-headed* and I inform you that the outcome of the coin toss is heads, how much information have I given you.
(d) Which of the six coins has the most uncertain coin toss outcome? Which has the least uncertain outcome?
(e) I toss one of the six coins and inform you of the outcome. For which coin will I have given you the most information.

---

111 Definition (Uncertainty/Information)
A function $I(E)$ that gives the uncertainty/information of an event E should have the following properties:
(i) $I(E) \geq 0$
(ii) If $P(E_1) \leq P(E_2)$, then $I(E_1) \geq I(E_2)$.
(iii) If E_1 and E_2 are independent events, i.e. $P(E_1 \cap E_2) = P(E_1)P(E_2)$, then $I(E_1 \cap E_2) = I(E_1) + I(E_2)$.

---

112 Definition (Self-Information)
Assume $P(E) \neq 0$. The self-information associated with event E is given by

$$I(E) = -\log_2(P(E))$$

where $I(E)$ has units of *bits* of information.

---

113 Example (Fair Coin Toss)
If I tell you that the outcome of a fair coin toss is heads, how much information have I given you?
