CMSC 858T: Randomized Algorithms

Spring 2003

Handout 2: Some relevant bounds

We discuss some simple bounds that are useful in the design and analysis of randomized algorothms. Throughout this handout, log denotes log base 2, and ln denotes the natural logarithm (i.e., log base e).

The binomial coefficients. The binomial coefficients $\binom{n}{r}$ pop up frequently, and here are some useful related bounds. First, you may know of Stirling's approximation: $r! \sim \sqrt{2\pi r}(r/e)^r$. A more precise result is due to Robbins [1]:

$$(r/e)^r \sqrt{2\pi r} e^{1/(12r+1)} \leq r! \leq (r/e)^r \sqrt{2\pi r} e^{1/(12r)},$$

valid for all $r \geq 1$. Using this, show that $\binom{n}{\lfloor n/2 \rfloor} = (\sqrt{2/\pi} + o(1)) \cdot 2^n/\sqrt{n}$, where the "$o(1)$" term is a function of n that tends to zero as $n \to \infty$. Next, using the fact that $\binom{n}{r} = n(n-1)(n-2)\cdots(n-r+1)/r!$, show the following useful bounds:

$$(n/r)^r \leq \binom{n}{r} \leq (ne/r)^r. \tag{1}$$

In fact, there is a sharper version of the second inequality in (1); it is in fact true that

$$\sum_{i=0}^{r} \binom{n}{i} \leq (ne/r)^r. \tag{2}$$

The above bounds are usually sufficient when $r \ll n$. The other important case is when $r = \Theta(n)$, i.e., where r is quite large. The *binary entropy function* $H(p) = -p \log p - (1-p)\log(1-p)$ where $p \in [0,1]$, will play a crucial role here. (When $p = 0$ or 1, we take the limiting value of $H(p)$; thus, we define $H(0) = H(1) = 0$.) Since $\binom{n}{i} = \binom{n}{n-i}$, let us assume that $r \leq n/2$. Now, for some constant $c \leq 1/2$, consider the case where $r = cn$. Using Robbins' formula, convince yourself that in this case,

$$\binom{n}{r} = 2^{(H(c)-o(1))\cdot n}, \tag{3}$$

where the "$o(1)$" is a function of n that tends to zero as $n \to \infty$. Thus, we have a fairly precise estimate of $\binom{n}{r}$ in the case where $r = \Theta(n)$. It is also true that

$$\sum_{i=0}^{r} \binom{n}{i} \leq 2^{H(c)\cdot n}. \tag{4}$$

Inequalities via basic calculus. One basic inequality we will use frequently is that $1 + x \leq e^x$ for all real x, with equality iff $x = 0$. One can prove such inequalities by elementary calculus, as follows. Define $f(x) = e^x - 1 - x$. We have $f(0) = 0$, and $f'(x) = e^x - 1$; thus, $f'(x)$ is positive for $x > 0$, and negative for $x < 0$. Therefore, f is increasing for $x > 0$, and decreasing for $x < 0$. This means that $f(x) > f(0)$ for all nonzero x; thus, $f(x)$ is positive for all nonzero x.

Such a simple approach using derivatives is often made easier by first taking logarithms on both sides of an inequality that we want to prove. Suppose we want to show that $a \leq b$, for some positive quantities a and b. Since $\ln x$ is a monotone increasing function of x for $x > 0$, proving $a \leq b$ is equivalent to proving $\ln a \leq \ln b$, which is sometimes useful. As an example, suppose you want to show that $1 + x >= e^{x - x^2/2}$ for all $x \geq 0$; this is equivalent to showing that $\ln(1+x) \geq x - x^2/2$ for all $x \geq 0$. Now define $g(x) = \ln(1+x) - x + x^2/2$. Observe that $g(0) = 0$ and that $g(x) = 1/(1+x) - 1 + x = x^2/(1+x)$, which is positive for $x > 0$. This completes the proof.

The AM-GM inequality. Another useful inequality to keep in mind is the arithmetic mean – geometric mean inequality: if we have non-negative reals $a_1, a_2, \ldots, a_n \geq 0$, then the arithmetic mean (i.e., average) of these numbers is at least as large as their geometric mean:

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \geq (a_1 a_2 \cdots a_n)^{1/n};$$

equality holds here iff all the a_i are equal.

Convexity. A function $f : \Re \to \Re$ is *convex* in the interval $[a, b]$ if and only if for all a' and b' such that $a \leq a' \leq b' \leq b$ and for all $q \in [0, 1]$,

$$f(qa' + (1-q)b') \leq qf(a') + (1-q)f(b'). \tag{5}$$

In words, for all $[a', b'] \subseteq [a, b]$, we have the following: in the domain $x \in [a', b']$, the function $f(x)$ lies below the straight line joining the points $(a', f(a'))$ and $(b', f(b'))$. f is called concave in $[a, b]$ if the inequality "$\leq$" in (5) is replaced by a "$\geq$"; in particular, f is concave iff the function $-f$ is convex.

A very useful way to check for convexity is as follows: if the second derivative $f''(x)$ exists in the interval $[a, b]$, then: (i) f is convex in $[a, b]$ iff $f''(x) \geq 0$ for all $[a, b]$, and (ii) f is concave in $[a, b]$ iff $f''(x) \leq 0$ for all $[a, b]$. Thus, for instance, it is easy to check that functions such as x^2, e^{tx} (for any constant t) are convex everywhere, while, for instance, $\ln x$ and $\sqrt{x}$ are concave in the region $x > 0$.

One useful application of convexity/concavity is as follows. Suppose a function f is convex. Then, suppose we have variables $x_1, x_2, \ldots, x_n$ subject to the constraint $\sum_i x_i = a$, for some constant a. Then, $\sum_i f(x_i)$ is minimized when all the x_i are equal; i.e., equal to a/n. The same remark holds if f is concave, and if we wish to maximize $\sum_i f(x_i)$. Thus, for instance, we have for any reals $x_1, x_2, \ldots, x_n$ and any real t, that:

$$\sum_{i=1}^{n} x_i^2 \geq (\sum_i x_i)^2/n; \quad \sum_{i=1}^{n} e^{tx_i} \geq ne^{t(x_1+x_2+\cdots+x_n)/n};$$

prove these inequalities.

The complementary remarks hold for convex maximization and concave minimization. Suppose we have variables $x_1, x_2, \ldots, x_n$ subject to the constraints $\sum_i x_i = a$, and such that $x_i \in [b, c]$ for all i; a, b and c are some constants here. Then, if f is convex in the interval $[b, c]$, $\sum_i f(x_i)$ is maximized when all the x_i but for at most one, lie in the set $\{b, c\}$. The same remark holds if f is concave, and if we wish to minimize $\sum_i f(x_i)$.

References

[1] H. Robbins. A remark on Stirling's formula. *Amer. Math. Monthly*, 62:26–29, 1955.