Numbers and Numeric Computations

1 Computing $f(x) = \sin x$

1.1 Why Worry?

We do not really need to worry about computing $\sin x$, since all programming languages used by engineers, including MATLAB, have it built in. It is, nonetheless, instructive thinking about how this computation is accomplished.

1.2 Series Expansion

The *Maclaurin Series*[1] for $\sin x$ is

$$\sin x = \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)!} x^{2j+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \tag{1}$$

This is easily coded, but for the infinite upper limit on the sum. So, we truncate the sum after "enough terms".

$$\sin(x) \approx \sum_{j=0}^{n-1} \frac{(-1)^j}{(2j+1)!} x^{2j+1} \tag{2}$$

Ok, how many terms is "enough"?

1.3 Error Term

The right-hand-side of (1) is a convergent, alternating series. From calculus we know that the error introduced by approximating an alternating series by its partial sum is less than the absolute value of the first truncated term[2]:

$$\text{error}(x, n) = \left| \sum_{j=n}^{\infty} \frac{(-1)^j}{(2j+1)!} x^{2j+1} \right| \leqslant \left| \frac{(-1)^n}{(2n+1)!} x^{2n+1} \right| = \frac{|x|^{2n+1}}{(2n+1)!} \tag{3}$$

Let's investigate this error term.

$$\frac{|x|^{2n+1}}{(2n+1)!} = \frac{|x|}{1} \cdot \frac{|x|}{2} \cdot \frac{|x|}{3} \cdots \frac{|x|}{(2n+1)} \tag{4}$$

For any fixed value of x we can choose a value of n such that $(2n + 1) \gg |x|$. More to the point, for any fixed value of x we can choose an n such that the error term, (3), is less than some specified tolerance. But, the required number of terms is an increasing function of $|x|$: the bigger $|x|$ is, the bigger n must be.

[1] You do not need to have a multitude of series expansions memorized. Nonetheless, you do need to (a) know that they exist, (b) know where to find them, and (c) know how to use them. See, for example, `http://en.wikipedia.org/wiki/Taylor_series`.

[2] This is a useful fact to keep in your mind if you are doing numeric computing. I drag it out once or twice a year for a drive, so I remember it. See `http://en.wikipedia.org/wiki/Alternating_series` for more information.

1.4 Escaping to a Friendly Place

The sin has many interesting symmetries. For example, we know[3] that

- $\sin(x) = -\sin(-x)$
- $\sin(x) = \sin(x \pm 2\pi) = \sin(x \pm 4\pi) = \ldots$
- $\sin(x) = -\sin(x + \mathtt{pi})$
- $\sin(x) = \sin(\pi - x)$

Using these symmetries we can take any argument x and map it to an equivalent value in a small range. Let

$$N = \mathtt{round}\left(\frac{|x|}{\pi}\right) \tag{5}$$

and

$$y = |x| - N\pi \tag{6}$$

Then

$$|x| = N\pi + y \tag{7}$$

with

$$|y| \leqslant \frac{\pi}{2} \tag{8}$$

and finally,

$$\boxed{\sin(x) = \sin(y) \cdot (-1)^N \cdot \mathtt{sign}(x)} \tag{9}$$

With this bit of fancy footwork, we have transformed our problem from computing the $\sin(x)$ over an infinite domain, to computing the $\sin(y)$, over a very limited domain.

1.5 How Many Terms are Enough?

According to (3), our error introduce when computing $\sin(y)$ using a partial sum is bounded by

$$\mathrm{error}(y, n) \leqslant \frac{|y|^{2n+1}}{(2n+1)!} \leqslant \frac{(\pi/2)^{2n+1}}{(2n+1)!} \tag{10}$$

Since

$$\sin(\pi/2) = 1 \tag{11}$$

any error greater than $\mathtt{eps}$ will be lost by truncation, we can compute our necessary number of terms using

$$\frac{(\pi/2)^{2n+1}}{(2n+1)!} \leqslant \mathtt{eps} \tag{12}$$

Doing a bit of computing we find

$$\frac{(\pi/2)^{2\cdot(11)+1}}{(2\cdot(11)+1)!} \leqslant \mathtt{eps} \leqslant \frac{(\pi/2)^{2\cdot(10)+1}}{(2\cdot(10)+1)!} \tag{13}$$

From which we conclude that 11 terms is always enough.

[3] By "know" I do not mean to imply that we should have these identities memorized. Rather, I mean that with a simple sketch we can figure each of these out, or verify them when someone reminds us that they exist.

1.6 Final Form

Putting the piece together we have

$$\sin(y) \approx y - \frac{y^3}{3!} + \frac{y^5}{5!} - \frac{y^7}{7!} + \frac{y^9}{9!} - \frac{y^{11}}{11!} + \frac{y^{13}}{13!} - \frac{y^{15}}{15!} + \frac{y^{17}}{17!} - \frac{y^{19}}{19!} + \frac{y^{21}}{21!} - \frac{y^{23}}{23!} \tag{14}$$

```
function [z] = mysin(x)

N = round(abs(x)/pi);
y = abs(x) - N*pi;

term = y;
sum  = y;
for j = 1:11
    term = -term * y*y/(2*j)/(2*j + 1);
    sum = sum + term;
end

if bitget(N,1)
    z = -sign(x) * sum;
else
    z =  sign(x) * sum;
end
```

2 Computing $f(x) = \frac{\sin x}{x}$

2.1 Why Worry?

In general, we compute $\frac{\sin x}{x}$ by computing $\sin x$ and dividing it by x. Yes,

$$f(x) = \frac{\sin x}{x} \tag{15}$$

done. However, as $x \to 0$ we find the numerator and denominator going to 0, and 0/0 gives a computer indigestion. This is an *indeterminant form*[4].

2.2 L'Hospital's Rule

We apply our friend L'Hospital's Rule[5] and find

$$\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \frac{\frac{d \sin x}{dx}}{\frac{dx}{dx}} = \lim_{x\to 0} \frac{\cos x}{1} = 1 \tag{16}$$

Thus, we have a "removable singularity" at $x = 0$. Our function should return 1 for $\frac{\sin x}{x}$ evaluated at $x = 0$, not blow up.

[4]See, for example, `http://en.wikipedia.org/wiki/Indeterminate_form`.

[5]See, for example, `http://en.wikipedia.org/wiki/L'Hopital's_rule`.

2.3 Series Expansion

We replace $\sin x$ by its *Maclaurin Series* (1), and write

$$\frac{\sin x}{x} = \frac{x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots}{x} \tag{17}$$

Dividing the numerator and denominator by x yields[6]

$$\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \cdots \tag{18}$$

Unlike the previous section, where we investigated $\sin x$, $\frac{\sin x}{x}$ does not have all of the nice symmetries. Fortunately, we do not need to worry about large x, since when x is large we simply compute $\sin x$ and dividing it by x.

When x is close to 0, however, we can use the series expansion (18) to generate our approximation. Say

$$\frac{\sin x}{x} \approx 1 - \frac{x^2}{3!} \tag{19}$$

Yes, this is an arbitrary choice.

2.4 Asking the Right Question

As mentioned above, the problems with $\frac{\sin x}{x}$ are not the same as the problems with $\sin x$. In this case we need to decide when to switch from the obvious evaluation mode, (15), to a special computation in the neighborhood of $x = 0$. In this case, we choose the n and compute the value of x at which we switch.

2.5 Error Term

As before, the right-hand-side of (18) is a convergent, alternating series. From calculus we know that the error introduced by approximating an alternating series by its partial sum is less than the absolute value of the first truncated term:

$$\text{error}(x) = \left| \sum_{j=2}^{\infty} \frac{(-1)^j}{(2j+1)!} x^{2j} \right| \leqslant \left| -\frac{x^4}{5!} \right| = \frac{|x|^4}{5!} \tag{20}$$

Since $\frac{\sin x}{x}$ is on the order of 1 in the neighborhood of $x = 0$, any error less than `eps` gets lost in the truncation. So, we can reasonably use `eps` as our error threshold.

$$\frac{|x|^4}{5!} \leqslant \texttt{eps} \tag{21}$$

which yields

$$|x| \leqslant (120 \cdot \texttt{eps})^{1/4} \approx 4.0402 \times 10^{-4} \tag{22}$$

[6]From (18) we can see immediately that

$$\lim_{x \to 0} \frac{\sin x}{x} = 1$$

2.6 Final Form

```
function [z] = fred(x)

if abs(x) <  4.0402e-004
    z = 1 - x*x/6;
else
    z = sin(x)/x;
end
```

3 Computing $f(x) = \exp(x)$

3.1 Why Worry?

As with $\sin x$, we do not really need to worry about computing $\exp(x)$, since all programming languages used by engineers, including MATLAB, have it built in. It is, again, instructive thinking about how this computation is accomplished.

3.2 Series Expansion

The *Maclaurin Series* for $\exp(x)$ is

$$\exp(x) = \sum_{j=0}^{\infty} \frac{x^j}{j!} \tag{23}$$

This is easily coded, but for the infinite upper limit on the sum. So, we truncate the sum after "enough terms".

$$\exp(x) \approx \sum_{j=0}^{n-1} \frac{x^j}{j!} \tag{24}$$

Ok, how many terms is "enough"? Interestingly, lots and lots if you do the computation naively.

3.3 Escaping to a Friendly Place

Let

$$N = \texttt{round}\left(\frac{x}{\log 2}\right) \tag{25}$$

and

$$y = x - N \log 2 \tag{26}$$

Then we can write

$$x = N \log 2 + y \tag{27}$$

where N is an integer, and

$$|y| \leqslant \frac{\log 2}{2} \approx 0.3466 \tag{28}$$

Putting the pieces together, we find

$$\exp(x) = \exp(N \log 2 + y) = \exp(N \log 2) \cdot \exp(y) = 2^N \cdot \exp(y) \tag{29}$$

The advantage of this is that the term 2^N requires nothing more than changing the exponent part of the floating point number, and $\exp(y)$ over the small domain (i.e. -0.3466 to $+0.3466$) can be computed accurately and efficiently using a Pade approximation.

3.4 Final Form

This is an example of a Pade approximation[7] formulation for $\exp(x)$.

```
function z = myexp(y)

Num = ((((y + 30)*y + 420)*y + 3360)*y + 15120)*y + 30240;
Den = ((((y - 30)*y + 420)*y - 3360)*y + 15120)*y - 30240;
z = -Num/Den;
```

[7] See, for example, http://en.wikipedia.org/wiki/Pade_approximant and http://en.wikipedia.org/wiki/Pade_table.