~ Key ~

Math 701 Quiz 1 Version A

1. Find a suitable trigonometric identity so that $1 - \cos x$ can be accurately computed for small x with calls to the system functions for $\sin x$ or $\cos x$.

$$\sin(a+b) = \sin a \cos b + \cos a \sin b$$
$$\cos(a+b) = \cos a \cos b - \sin a \sin b$$
$$\cos x = \cos\left(\frac{x}{2} + \frac{x}{2}\right) = \cos^2\frac{x}{2} - \sin^2\frac{x}{2}$$
$$-\cos x = \sin^2\frac{x}{2} - \cos^2\frac{x}{2}$$
$$1 = \sin^2\frac{x}{2} + \cos^2\frac{x}{2}$$

$$1 - \cos x = 2\sin^2\frac{x}{2}$$

2. State Taylor's theorem including all hypothesis and the remainder term.

Let $f:\mathbb{R}\to\mathbb{R}$ be a $n+1$ times continuously differentiable function.

Then
$$f(x+h) = \sum_{k=0}^{n} \frac{h^n}{n!} f^{(n)}(x) + \frac{h^{n+1}}{(n+1)!} f^{(n+1)}(\xi)$$

for some ξ between x and $x+h$.

3. State Newton's method.

Let $f:\mathbb{C}\to\mathbb{C}$ be a differentiable such that $f(p)=0$. Newton's method forms a sequence of approximations of p by starting with an initial approximation x_0 and defining
$$x_{n+1} = g(x_n) \quad \text{where} \quad g(x) = x - f(x)/f'(x).$$

4. Prove only one of the following:

(i) Taylor's theorem.

(ii) Let f be a twice continuously differentiable function and p be a point such that $f(p) = 0$ and $f'(p) \neq 0$. Prove that Newton's method is quadratically convergent provided x_0 is close enough to p.

(i) Taylor's Theorem: By the Fundamental Theorem of Calculus

$$f(x+h) - f(x) = \int_x^{x+h} f'(t)dt.$$

Now integrate by parts taking $u = f'(t)$, $dv = dt$

$du = f''(t)dt \quad v = t + c$

Thus

$$\int_x^{x+h} f'(t)dt = (t+c)f'(t)\Big|_x^{x+h} - \int_x^{x+h} (t+c)f''(t)dt$$

Choose the constant $c = -x-h$ so plugging $t = x+h$ in the first term on the right hand side of the above gives zero.

$$\int_x^{x+h} f'(t) = hf'(x) - \int_x^{x+h} (t-x-h)f''(t)dt$$

$$= hf'(x) + \int_x^{x+h} (x+h-t)f''(t)dt$$

Integrate by parts again using $u = f''(t)$ $dv = x+h-t$

to obtain $\quad du = f'''(t)dt \quad v = -\frac{(x+h-t)^2}{2}$

$$\int_x^{x+h} (x+h-t)f''(t)dt = -\frac{(x+h-t)^2}{2} f''(t)\Big|_x^{x+h} + \int_x^{x+h} \frac{(x+h-t)^2}{2} f'''(t)dt$$

$$= \frac{h^2}{2} f''(x) + \int_x^{x+h} \frac{(x+h-t)^2}{2} f'''(t)dt$$

Repeatedly integrate by parts until arriving at the $n+1$ derivative to obtain

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2} f''(x) + \cdots + \frac{h^n}{n!} f^{(n)}(x) + \int_x^{x+h} \frac{(x+h-t)^n}{n!} f^{(n+1)}(t)dt.$$

Math 701 Quiz 1 Version A

Proof of Taylor's theorem or the quadratic convergence of Newton's method continues ...

Taylor's Theorem continues...

Now apply the weighted mean-value theorem for integrals to the last term taking the weights

$$w(t) = \frac{(x+h-t)^n}{n!} \text{ for } h \geq 0 \text{ or } \left(\frac{t-x-h}{n!}\right)^n \text{ for } h<0.$$

Thus, when $h \geq 0$ we have

$$\int_x^{x+h} w(t) f^{(n+1)}(t)\,dt = \left(\int_x^{x+h} w(t)\,dt\right) f^{(n+1)}(\xi) \text{ for some } \xi \in [x, x+h]$$

Since $\displaystyle\int_x^{x+h} w(t)\,dt = \int_x^{x+h} \frac{(x+h-t)^n}{n!}\,dt = \frac{h^{n+1}}{(n+1)!}$ the result follows.

If $h<0$ then for some $\xi \in [x+h, x]$ we have

$$\int_{x+h}^{x} w(t) f^{(n+1)}(t) = \left(\int_{x+h}^{x} \frac{(t-x-h)^n}{n!}\,dt\right) f^{(n+1)}(\xi) = \frac{(-h)^{n+1}}{(n+1)!} f^{(n+1)}(\xi).$$

Consequently

$$\int_x^{x+h} \frac{(x+h-t)^n}{n!} f^{(n+1)}(t) = (-1)^{n+1} \int_{x+h}^{x} \frac{(t-x-h)^n}{n!} f^{n+1}(t)\,dt$$

$$= (-1)^{n+1} \int_{x+h}^{x} w(t) f^{(n+1)}(t) = \frac{h^{n+1}}{(n+1)!} f^{(n+1)}(\xi)$$

and again the result follows.

4. (ii) Prove the quadratic convergence of Newtons method.

By Taylor's theorem there are ξ_n between x_n and p such that

$$f(p) = f(x_n) + (p-x_n)f'(x_n) + \frac{(p-x_n)^2}{2} f''(\xi_n).$$

Let $e_n = p - x_n$. Then

$$e_{n+1} = p - x_{n+1} = p - g(x_n) = p - \left(x_n - \frac{f(x_n)}{f'(x_n)}\right) = e_n + \frac{f(x_n)}{f'(x_n)}.$$

assuming x_n is close enough to p then $f'(p) \neq 0$ implies that $f'(x_n) \neq 0$. Dividing Taylor's theorem by $f'(x_n)$ and using the fact that $f(p)=0$ now obtains

$$0 = \frac{f(x_n)}{f'(x_n)} + e_n + \frac{e_n^2}{2}\,\frac{f''(\xi_n)}{f'(x_n)}.$$

Consequently

$$e_{n+1} = -\frac{e_n^2}{2}\,\frac{f''(\xi_n)}{f'(x_n)}$$

Now let $\delta > 0$ be choosen so that $|f'(t)| > \frac{1}{2}|f'(p)|$ for all $|t-p| < \delta$ and define $A = \max\{|f''(t)| : |t-p| \leq \delta\}$. Note that since f'' is continuous this maximum exists. It follows that

$$|p - x_{n+1}| \leq M|p-x_n|^2 \quad \text{for } M = \frac{A}{|f'(p)|}.$$

for every x_n such that $|x_n - p| < \delta$. assuming x_n converges to p this shows quadratic convergence. To guarantee that x_n converges it is sufficient to choose x_0 close enough to p such that $M|p-x_0| < 1$.

Math 701 Quiz 1 Version A

5. Let $A \in \mathbf{R}^{d\times d}$ be a symmetric positive semidefinite matrix. Consider the

> **Power Method.** Choose $x_0 \in \mathbf{R}^d$ randomly. Then recursively compute $y_n = Ax_n$ and $x_{n+1} = y_n/\|y_n\|$ for $n \geq 0$.

Show for almost every choice of x_0 that the limits

$$\lambda = \lim_{n\to\infty} \|y_n\| \qquad \text{and} \qquad \xi = \lim_{n\to\infty} x_n$$

exist and that λ and ξ form an eigenvalue-eigenvector pair for A such that $A\xi = \lambda\xi$.

Since A is symmetric the spectral theorem for symmetric matrices implies there is an orthonormal basis of eigenvectors ξ_k and eigenvalues λ_k such that $A\xi_k = \lambda_k \xi_k$ and $\xi_i \cdot \xi_j = \begin{cases} 1 & \text{for } i=j \\ 0 & \text{otherwise.} \end{cases}$

Moreover, since A is positive semidefinite then $\lambda_k \geq 0$ for all k. Further assume $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$. Given $x_0 \in \mathbb{R}^d$ choose c_k such that

$$x_0 = c_1\xi_1 + c_2\xi_2 + \cdots + c_d\xi_d.$$

Since x_0 was chosen randomly, then with probability one it must happen that $c_1 \neq 0$. By definition

$$y_0 = Ax_0 = c_1\lambda_1\xi_1 + c_2\lambda_2\xi_2 + \cdots + c_d\lambda_d\xi_d.$$

and furthermore

$$x_1 = \frac{c_1\lambda_1\xi_1 + c_2\lambda_2\xi_2 + \cdots + c_d\lambda_d\xi_d}{\|c_1\lambda_1\xi_1 + c_2\lambda_2\xi_2 + \cdots + c_d\lambda_d\xi_d\|}$$

By induction it follows that

$$x_n = \frac{c_1\lambda_1^n\xi_1 + c_2\lambda_2^n\xi_2 + \cdots + c_d\lambda_d^n\xi_d}{\|c_1\lambda_1^n\xi_1 + c_2\lambda_2^n\xi_2 + \cdots + c_d\lambda_d^n\xi_d\|}$$

and

$$y_n = \frac{c_1\lambda_1^{n+1}\xi_1 + c_2\lambda_2^{n+1}\xi_2 + \cdots + c_d\lambda_d^{n+1}\xi_d}{\|c_1\lambda_1^n\xi_1 + c_2\lambda_2^n\xi_2 + \cdots + c_d\lambda_d^n\xi_d\|}$$

Math 701 Quiz 1 Version A

Proof of the convergence of the power method continues ...

Now choose q so that

$$\lambda_1 = \lambda_2 = \cdots = \lambda_q \geq \lambda_{q+1} \geq \cdots \geq \lambda_d .$$

Then $\left(\frac{\lambda_k}{\lambda_1}\right)^n \to 0$ as $n \to \infty$ for $k > q$

and $\left(\frac{\lambda_k}{\lambda_1}\right) = 1$ for $k \leq q$.

Consequently

$$\|y_n\| = \frac{\|c_1\lambda_1^{n+1}\xi_1 + \cdots + c_q\lambda_q^{n+1}\xi_q + c_{q+1}\lambda_{q+1}^{n+1}\xi_{q+1} + \cdots + c_d\lambda_d^{n+1}\xi_d\| / \lambda_1^n}{\|c_1\lambda_1^n\xi_1 + \cdots + c_q\lambda_q^n\xi_q + c_{q+1}\lambda_{q+1}^{n+1}\xi_{q+1} + \cdots + c_d\lambda_d^n\xi_d\| / \lambda_1^n}$$

$$= \frac{\|(c_1\xi_1 + \cdots + c_q\xi_q)\lambda_1 + c_{q+1}\lambda_{q+1}\left(\frac{\lambda_{q+1}}{\lambda_1}\right)^n\xi_{q+1} + \cdots + c_d\lambda_d\left(\frac{\lambda_d}{\lambda_1}\right)^n\xi_d\|}{\| c_1\xi_1 + \cdots + c_q\xi_q + c_{q+1}\left(\frac{\lambda_{q+1}}{\lambda_1}\right)^n\xi_{q+1} + \cdots + c_d\left(\frac{\lambda_d}{\lambda_1}\right)^n\xi_d\|}$$

$$\to \frac{\|(c_1\xi_1 + \cdots + c_q\xi_q)\lambda_1\|}{\|c_1\xi_1 + \cdots c_q\xi_q\|} = |\lambda_1| = \lambda_1 \quad \text{as } n \to \infty$$

This shows that $\lim_{n\to\infty} \|y_n\|$ is an eigenvalue.

Proof of convergence of the power method continues...

To finish the proof note that, for similar reasons

$$x_n \to \frac{c_1\xi_1 + \cdots + c_q\xi_q}{\|c_1\xi_1 + \cdots + c_q\xi_q\|} \quad \text{as} \quad n \to \infty.$$

Now, since $\lambda_1 = \lambda_2 = \cdots = \lambda_q$ then

$$A\left(\frac{c_1\xi_1 + \cdots + c_q\xi_q}{\|c_1\xi_1 + \cdots + c_q\xi_q\|}\right) = \frac{c_1\lambda_1\xi_1 + \cdots + c_q\lambda_q\xi_q}{\|c_1\xi_1 + \cdots + c_q\xi_q\|}$$

$$= \frac{c_1\lambda_1\xi_1 + \cdots + c_q\lambda_1\xi_q}{\|c_1\xi_1 + \cdots + c_q\xi_q\|} = \lambda_1 \frac{c_1\xi_1 + \cdots + c_q\xi_q}{\|c_1\xi_1 + \cdots + c_q\xi_q\|}$$

Shows $\lim_{n\to\infty} x_n$ is an eigenvector.

Math 701 Quiz 1 Version A

6. For $x \in \mathbf{R}^d$ define $\|x\|_p = \Big(\sum_{k=1}^{d} |x_k|^p\Big)^{1/p}$.

(i) Prove that $\|x\|_2 \le \|x\|_1$.

(ii) [Extra credit] Prove or disprove that $\|x\|_p \le \|x\|_1$ for every $p \ge 1$.

By definition $\|x\|_2^2 = \sum_{k=1}^{d} |x_k|^2$

and $\|x\|_1^2 = \Big(\sum_{k=1}^{d} |x_k|\Big)^2 = \sum_{k=1}^{d} |x_k|^2 + 2\sum_{k>l} |x_k||x_l|$

since $2\sum_{k>l} |x_k||x_l| \geqslant 0$ it follows that $\|x\|_2^2 \le \|x\|_1^2$.

Taking square roots now yields that $\|x\|_2 \le \|x\|_1$.

Extra credit: Upon taking powers of both sides we equivalently need to show that

$$\sum_{k=1}^{d} |x_k|^p \le \left(\sum_{k=1}^{d} |x_k| \right)^p$$

We procede be induction on d. If $d=1$ the result is obvious. Suppose the result is true for $d-1$. Then

$$\sum_{k=1}^{d} |x_k|^p = \sum_{k=1}^{d-1} |x_k|^p + |x_d|^p \le \left(\sum_{k=1}^{d-1} |x_k| \right)^p + |x_d|^p$$

Define $a = \sum_{k=1}^{d-1} |x_k|$ and $b = |x_d|$.

Claim that $a^p + b^p \le (a+b)^p$. If $a=0$ this is obvious. Consequently we may assume $a \ne 0$. Then writing $a+b = a(1+x)$ where $x = b/a$. It follows that proving $1+x^p \le (1+x)^p$ is sufficient. To show this define $f(x) = (1+x)^p - 1 - x^p$. Then

$$f'(x) = p(1+x)^{p-1} - p x^{p-1} \ge 0 \text{ for } x \ge 0$$

implies f is increasing. Since $f(0) = 0$ and f is increasing then $f(x) \ge 0$ for all $x \ge 0$. Consequently we have $1 + x^p \le (1+x)^p$ for all $x \ge 0$. Thus

$$\sum_{k=1}^{d} |x_k|^p \le a^p + b^p = a^p(1+x^p) \le a^p(1+x^p) = (a+b)^p = \left(\sum_{k=1}^{d} |x_k| \right)^p$$

finishes the induction and the proof.