MSCF Mathematics Preparatory Course
August 2006
Solutions to Homework #4 Exercises

7.1 (b) First we derive the Taylor series centered at 0 associated with f. We note that

$$\begin{aligned} f^{(0)}(x) &= f(x) = \sin x \\ f'(x) &= \cos x \\ f''(x) &= -\sin x \\ f'''(x) &= -\cos x \\ f^{(4)}(x) &= \sin x \end{aligned}$$

and now we see that taking higher derivatives will result in cycling through the same four functions. Evaluating f and its derivatives at 0 generates the values

$$0, 1, 0, -1, 0, 1, 0, -1, \ldots.$$

So the Taylor coefficients are as follows:

$$\begin{aligned} &\frac{0}{0!}, \frac{1}{1!}, \frac{0}{2!}, \frac{-1}{3!}, \frac{0}{4!}, \frac{1}{5!}, \frac{0}{6!}, \frac{-1}{7!}, \ldots \\ &= 0, 1, 0, -1/3!, 0, 1/5!, 0, -1/7!, \ldots. \end{aligned}$$

The Taylor series $T(x)$ associated with f (and centered at 0) is therefore

$$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \ldots.$$

Since only odd-order terms are present, we can write this series concisely by observing that for each $i = 0, 1, 2, 3, \ldots$, the number $2i + 1$ is an odd integer, and the formula $2i + 1$ generates the sequence $1, 3, 5, 7, \ldots$. To get the alternating signs in the Taylor coefficients, we can introduce the factor $(-1)^i$. So we have

$$T(x) = \sum_{i=0}^{\infty} \frac{(-1)^i x^{2i+1}}{(2i+1)!}.$$

Though you were not asked to do so, I will show that $f(x) = T(x)$ at least for x in some interval $(-r, r)$, $r > 0$, so that Proposition 7.5 yields the analyticity of f at 0. Let us fix $x \in \mathbf{R}$. Using (33), we have

$$\sin x - T_N(x) = \frac{f^{(N+1)}(z)}{(N+1)!} \cdot x^{N+1},$$

for some z between 0 and x. We must show that the limit of the right hand side is 0 as N tends to infinity; whether this is true may depend on how close x is to 0. Now since $f^{(N+1)}$ is one of the functions $\sin(\cdot)$, $\cos(\cdot)$, $-\sin(\cdot)$, $-\cos(\cdot)$, we have that $|f^{(N+1)}(z)| \leq 1$, regardless of the location of z. Hence

$$\left| \frac{f^{(N+1)}(z)}{(N+1)!} \cdot x^{N+1} \right| \leq \frac{|x|^{N+1}}{(N+1)!},$$

and if

$$\lim_{N\to\infty} \frac{|x|^{N+1}}{(N+1)!} = 0,$$

then we are finished. But this is precisely (36), and we showed in the text (in the narrative following (36)), that this is true for any value of x. Hence f is analytic at 0, and moreover f agrees with its Taylor series on all of $\mathbf{R}$. Now we note that $P(x) = T_5(x)$, so given $x \in [-\pi, \pi]$, and applying Theorem 7.7 with $N = 5$, we have

$$\sin x = x - \frac{x^3}{6} + \frac{x^5}{120} + \frac{-\sin z}{7!} x^6,$$

where z is between 0 and x. Since $|-\sin z| \leq 1$, we have

$$|\sin x - T_5(x)| \leq \frac{|x|^6}{6!} \leq \frac{\pi^6}{6!} \approx 1.335262769.$$

This is a pretty crummy error bound, though, considering $-1 \leq \sin x \leq 1$. But perhaps we can improve the error estimate by noting that $P(x)$ is also equal to $T_6(x)$. Then we can apply Taylor's Theorem with $N = 6$ instead and obtain

$$\sin x = x - \frac{x^3}{6} + \frac{x^5}{120} + \frac{-\cos z}{7!} x^7,$$

or

$$f(x) - P(x) = \frac{-\cos z}{7!} x^7.$$

To find a uniform bound on the error associated with approximating $f(x)$ by $P(x)$, then, we note

$$|f(x) - P(x)| = \left| \frac{-\cos z}{7!} x^7 \right| \leq \frac{x^7}{7!},$$

since $|\cos z| \leq 1$. Then since $x \in [-\pi, \pi]$,

$$\frac{x^7}{7!} \leq \frac{\pi^7}{7!} \approx 0.599264529,$$

which is an improvement, but is still not very reassuring. Hence part (c).

(c) Our observations in part (b) tell us that given N,

$$|\sin x - T_N(x)| \leq \frac{|x|^{N+1}}{(N+1)!} \leq \frac{\pi^{N+1}}{(N+1)!}$$

holds for $x \in [-\pi, \pi]$, and also that we need only consider even values of N. So by calculator, we test $\pi^9/9! \approx 0.0821$, so that is not small enough. Next $\pi^{11}/11! \approx 0.00737 < 0.01$, and therefore $T_{10}(x)$, which is the same as

$$T_9(x) = x - \frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{5040} + \frac{x^9}{362,880},$$

will give the desired accuracy on $[-\pi, \pi]$. For reassurance, we calculate $T_9(5\pi/6) \approx 0.500949776$, and yes, we get two (in fact three) decimal places of accuracy.

7.2 (a) Since $f \in C^6(D)$, all partial derivatives of f up to and including order 6 are defined and continuous on D. In particular, $\dfrac{\partial^2 f}{\partial x_1 \partial x_5}$ exists and is continuous on D. By the theorem, then, $\dfrac{\partial^2 f}{\partial x_5 \partial x_1}$ also exists and is continuous, and moreover, for all $x \in D$,

$$\frac{\partial^2 f}{\partial x_1 \partial x_5}(x) = \frac{\partial^2 f}{\partial x_5 \partial x_1}(x).$$

Now since $\frac{\partial^2 f}{\partial x_1 \partial x_5}$ and $\frac{\partial^2 f}{\partial x_5 \partial x_1}$ represent the same function, we can consider the partial derivative of this function with respect to x_4, and we have

$$\frac{\partial^3 f}{\partial x_4 \partial x_1 \partial x_5} = \frac{\partial^3 f}{\partial x_4 \partial x_5 \partial x_1},$$

and now let us call this function g. Now $f \in C^6(D)$, which implies that $f \in C^4(D)$, and therefore $g \in C^1(D)$. We also note that $f \in C^5(D)$ implies that $\dfrac{\partial^2 g}{\partial x_3 \partial x_4}$ exists and is continuous on D. By the theorem, then, $\dfrac{\partial^2 g}{\partial x_4 \partial x_3}$ also exists and is continuous on D, and for all $x \in D$,

$$\frac{\partial^2 g}{\partial x_3 \partial x_4}(x) = \frac{\partial^2 g}{\partial x_4 \partial x_3}(x),$$

which translates to

$$\frac{\partial^5 f}{\partial x_3 \partial x_4^2 \partial x_5 \partial x_1}(x) = \frac{\partial^5 f}{\partial x_4 \partial x_3 \partial x_4 \partial x_1 \partial x_5}(x).$$

Denoting the above function by h, then, we have h representing two equal fifth-order partial derivatives of f. Since $f \in C^6(D)$, any first-order partial derivative of h exists; in particular $\frac{\partial h}{\partial x_3}$ exists, and

$$\frac{\partial h}{\partial x_3} \equiv \frac{\partial^6 f}{\partial x_3^2 \partial x_4^2 \partial x_5 \partial x_1} \equiv \frac{\partial^6 f}{\partial x_3 \partial x_4 \partial x_3 \partial x_4 \partial x_1 \partial x_5}, \qquad (**)$$

as desired.

(b) Note that in $(**)$ we have two "different", but identically equal, sixth-order partial derivatives of f which, in either case, involves differentiating with respect to x_1 once, with respect to x_3 twice, with respect to x_4 twice, and with respect to x_5 once. We may also note that we differentiate with respect to x_2 zero times. This suggests the multiindex $\alpha = (1, 0, 2, 2, 1)$, where the value of each α_i indicates how many times we differentiate with respect to the variable x_i. However, we need to know that the order in which the partial differentiations are performed is not an issue, and our argument in part (a) essentially shows this, for if we wish to consider any of the various sixth-order partial derivatives of f which are suggested by $\alpha = (1, 0, 2, 2, 1)$ (such as

$$\frac{\partial^6 f}{\partial x_5 x_4^2 x_3^2 x_1} \quad \text{or} \quad \frac{\partial^6 f}{\partial x_4 x_3^2 x_5 x_4 x_1} \quad \text{or} \quad \frac{\partial^6 f}{\partial x_1 x_3 x_5 x_4 x_3 x_4}),$$

we can show that any two are equal to each other, and to those in $(**)$, by applying the theorem as many times as necessary to "swap" consecutive differentiations. Thus we are justified in using the simplified notation

$$D^{(1,0,2,2,1)} f$$

to refer to any one of these sixth-order partial derivatives.

7.5 (a) Let $P(t, \omega)$ be the desired Taylor polynomial. Note that

$$f(t, \omega) = e^{\sigma\omega - \frac{1}{2}\sigma^2 t} = e^{\sigma\omega} \cdot e^{-\frac{1}{2}\sigma^2 t},$$

and each of the variables t and ω appear in only one of the factors on the right. Therefore, any partial derivative of f is a constant multiple of f; for example,

$$\frac{\partial f}{\partial t}(t, \omega) = -\frac{1}{2}\sigma^2 e^{\sigma\omega} \cdot e^{-\frac{1}{2}\sigma^2 t} = -\frac{1}{2}\sigma^2 f(t, \omega).$$

So for any partial derivative

$$\frac{\partial^k f}{\partial t^i \partial \omega^{k-i}}$$

of f, since $f(0,0) = 1$, we will have

$$\frac{\partial^k f}{\partial t^i \partial \omega^{k-i}}(0,0) = \left(-\frac{1}{2}\sigma^2\right)^i \cdot \sigma^{k-i} \cdot f(0,0) = \left(-\frac{1}{2}\sigma^2\right)^i \cdot \sigma^{k-i}.$$

So now we compute

$$\begin{array}{rcl} f(0,0) & = & 1 \\ \frac{\partial f}{\partial t}(0,0) & = & -\frac{1}{2}\sigma^2 \\ \frac{\partial f}{\partial \omega}(0,0) & = & \sigma \\ \frac{\partial^2 f}{\partial t^2}(0,0) & = & \frac{1}{4}\sigma^4 \\ \frac{\partial^2 f}{\partial t \partial \omega}(0,0) & = & -\frac{1}{2}\sigma^3 \\ \frac{\partial^2 f}{\partial \omega^2}(0,0) & = & \sigma^2 \end{array}$$

So now we can write down P:

$$P(t,\omega) = 1 - \frac{1}{2}\sigma^2 t + \sigma\omega + \frac{1}{8}\sigma^4 t^2 - \frac{1}{2}\sigma^3 t\omega + \frac{1}{2}\sigma^2\omega^2.$$

(b) Using Theorem 7.16 with $N = 2$, if (t,ω) is any point in the given set, then

$$f(t,\omega) = P(t,\omega) + \sum_{i=0}^{3} \frac{1}{i!(3-i)!}\left(\frac{\partial^3 f}{\partial t^i \partial \omega^{3-i}}(z_1,z_2)\right) \cdot t^i \omega^{3-i},$$

where (z_1, z_2) is some point on the line segment between $(0,0)$ and (t,ω). So we should start by determining the third order partial derivatives:

$$\begin{array}{rcl} \frac{\partial^3 f}{\partial t^3} & = & -\frac{1}{8}\sigma^6 \cdot f \\ \frac{\partial^3 f}{\partial t^2 \partial \omega} & = & \frac{1}{4}\sigma^5 \cdot f \\ \frac{\partial^3 f}{\partial t \partial \omega^2} & = & -\frac{1}{2}\sigma^4 \cdot f \\ \frac{\partial^3 f}{\partial \omega^3} & = & \sigma^3 \cdot f \end{array}$$

So the error in using P to approximate f at the point (t,ω) is

$$f(z_1,z_2)\left(-\frac{1}{8}\sigma^6 t^3 + \frac{1}{4}\sigma^5 t^2\omega - \frac{1}{2}\sigma^4 t\omega^2 + \sigma^3\omega^3\right). \quad (*)$$

Now note that since

$$f(t,\omega) = e^{\sigma\omega} \cdot e^{-\frac{1}{2}\sigma^2 t},$$

the factor on the left can be no larger than e^σ (when $\omega = 1$), and the factor on the right can be no larger than 1 (when $t = 0$). So $(*)$ is no larger than

$$e^\sigma \left|-\frac{1}{8}\sigma^6 t^3 + \frac{1}{4}\sigma^5 t^2\omega - \frac{1}{2}\sigma^4 t\omega^2 + \sigma^3\omega^3\right|$$

$$\le e^{\sigma}\left(\frac{1}{8}\sigma^6 t^3+\frac{1}{4}\sigma^5 t^2\omega+\frac{1}{2}\sigma^4 t\omega^2+\sigma^3\omega^3\right)$$
$$\le e^{\sigma}\left(\frac{1}{8}\sigma^6 T^3+\frac{1}{4}\sigma^5 T^2+\frac{1}{2}\sigma^4 T+\sigma^3\right),$$

where we have used the triangle inequality. So the last quantity is a uniform bound on the error associated with using P to approximate f on the given rectangle. Often this type of "rough" analysis, using the triangle inequality, is sufficient; finding the best possible upper bound on the error is usually very time-consuming and not worth the effort. However, by examining $(*)$ more closely, we can, in this case, improve the error bound considerably with not too much additional effort: Rewrite $(*)$ as

$$f(z_1,z_2)\cdot\frac{\sigma^3}{48}\left(8\omega^3-12\sigma\omega^2 t+6\sigma^2\omega t^2-\sigma^3 t^3\right)$$

and note that the last factor is equal to $(2\omega-\sigma t)^3$. As we have already noted that $f(z_1,z_2)$ can be no greater than e^{σ}, the error is no greater than

$$\frac{e^{\sigma}\sigma^3}{48}|2\omega-\sigma t|^3$$

Now we consider how large $|2\omega-\sigma t|^3$ may be on the given rectangle, and maximizing $|2\omega-\sigma t|^3$ is equivalent to maximizing $|2\omega-\sigma t|$. The key is to note that we have a linear expression; letting $g(t,\omega)=2\omega-\sigma t$, we observe that the *level sets* of g, that is, the curves in the plane on which g takes any given constant value, are straight lines. So for any $c\in\mathbf{R}$, the set $\{(t,\omega):g(t,\omega)=c\}$ is a line in the plane, with equation

$$2\omega-\sigma t=c,\qquad (**)$$

or in slope-intercept form,

$$\omega=\frac{\sigma}{2}t+\frac{c}{2}.$$

Now envision lines in the plane having slope $\sigma/2$ (which is a positive slope no greater than $1/2$). In view of $(**)$, maximizing $|2\omega-\sigma t|$ is tantamount to finding, among all lines intersecting

$$\{(t,\omega):0\le t\le T,0\le\omega\le 1\},$$

the one with ω-intercept (which is $c/2$) largest in magnitude. Draw a picture; the line which passes through $(0,0)$ is certainly not the one we seek, for then $c/2=0$, and by taking instead the line which intersects the rectangle at $(0,1)$, we get $c/2=1$. But note that if we consider the line intersecting the rectangle at $(T, 0)$, we get a line with ω-intercept equal to $-\sigma T/2$, which may be greater than 1 in magnitude; this of course depends on σ and T. So there are two possibilities: Either $|c|$ is maximized at $(0, 1)$ or at $(T, 0)$, so that

$$|2\omega - \sigma t| \leq \max\{2, \sigma T\}.$$

Hence we get the following simplified error bound:

$$|f(t,\omega) - P(t,\omega)| \leq \frac{e^{\sigma}\sigma^3}{48} \cdot \max\{8, \sigma^3 T^3\}$$

Problem A: We first observe that

$$\begin{aligned} x\exp\left\{\sigma\sqrt{\tau}\,y + \left(r - \frac{1}{2}\sigma^2\right)\tau\right\} \geq K \quad &\iff \quad \exp\left\{\sigma\sqrt{\tau}\,y + \left(r - \frac{1}{2}\sigma^2\right)\tau\right\} \geq \frac{K}{x} \\ &\iff \quad \sigma\sqrt{\tau}\,y + \left(r - \frac{1}{2}\sigma^2\right)\tau \geq \log\frac{K}{x} \\ &\iff \quad \log\frac{x}{K} + \left(r - \frac{1}{2}\sigma^2\right)\tau \geq -\sigma\sqrt{\tau}\,y \\ &\iff \quad d_-(\tau, x) \geq -y \\ &\iff \quad -d_-(\tau, x) \leq y. \end{aligned}$$

Therefore,

$$\begin{aligned} c(t,x) &= \int_{-d_-(\tau,x)}^{\infty} e^{-r\tau}\left(x\exp\left\{\sigma\sqrt{\tau}\,y + \left(r - \frac{1}{2}\sigma^2\right)\tau\right\} - K\right)\varphi(y)\,dy \\ &= \int_{-\infty}^{d_-(\tau,x)} e^{-r\tau}\left(x\exp\left\{-\sigma\sqrt{\tau}\,z + \left(r - \frac{1}{2}\sigma^2\right)\tau\right\} - K\right)\varphi(z)\,dz \\ &= x\int_{-\infty}^{d_-(\tau,x)} \exp\left\{-\sigma\sqrt{\tau}\,z - \frac{1}{2}\sigma^2\tau\right\}\varphi(z)\,dz - Ke^{-r\tau}\int_{-\infty}^{d_-(\tau,x)}\varphi(z)\,dz, \end{aligned}$$

where we have made the change of variable $z = -y$ and used the fact that $\varphi(-y) = \varphi(y)$. It remains to show that

$$\int_{-\infty}^{d_-(\tau,x)} \exp\left\{-\sigma\sqrt{\tau}\,z - \frac{1}{2}\sigma^2\tau\right\}\varphi(z)\,dz = N\big(d_+(\tau,x)\big), \tag{1}$$

$$\int_{-\infty}^{d_-(\tau,x)} \varphi(z)\,dz = N\big(d_-(\tau,x)\big). \tag{2}$$

Equation (2) follows immediately from the definition of $N(d_-(\tau, x))$. For equation (1), we make the second change of variable $w = z + \sigma\sqrt{\tau}$ to obtain

$$
\begin{aligned}
&\int_{-\infty}^{d_-(\tau,x)} \exp\left\{-\sigma\sqrt{\tau}\,z - \frac{1}{2}\sigma^2\tau\right\}\varphi(z)\,dz \\
&\quad = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_-(\tau,x)} \exp\left\{-\frac{1}{2}z^2 - \sigma\sqrt{\tau}\,z - \frac{1}{2}\sigma^2\tau\right\}\,dz \\
&\quad = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d-(\tau,x)} \exp\left\{-\frac{1}{2}(z+\sigma\sqrt{\tau})^2\right)\,dz \\
&\quad = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_+(\tau,x)} \exp\left\{-\frac{1}{2}w^2\right\}\,dw \\
&\quad = N\big(d_+(\tau,x)\big).
\end{aligned}
$$

Problem B: First we note a property of the dot product: If α is any scalar (that is, real number), and u and v are both vectors in any Cartesian space $\mathbf{R}^k$, then

$$\alpha(u \cdot v) = (\alpha u) \cdot v = u \cdot (\alpha v)$$

(I will leave it to you to verify this with straightforward computation.) Now recall from lecture that we have two different ways to view the product Ax: (i) as the vector whose components are the dot products of x with the rows of A, or (ii) as a linear combination of the columns of A, where the components of x are the scalar multiples. Following the notation used in lecture, we let $A^1, A^2, \ldots, A^m$ be the rows of A, and let $A_1, \ldots, A_n$ denote the columns of A. Then

$$(Ax) \cdot y = (x_1 A_1 + \ldots + x_n A_n) \cdot y$$

$$= (x_1 A_1) \cdot y + \ldots + (x_n A_n) \cdot y = x_1(A_1 \cdot y) + \ldots x_n(A_n \cdot y),$$

and this last quantity is the dot product of x with the vector

$$\begin{bmatrix} A_1 \cdot y \\ A_2 \cdot y \\ \cdot \\ \cdot \\ \cdot \\ A_n \cdot y \\ \end{bmatrix}$$

But as the columns of A are the rows of A^T, we may also write this vector as

$$\begin{bmatrix} (A^T)^1 \cdot y \\ (A^T)^2 \cdot y \\ \cdot \\ \cdot \\ \cdot \\ (A^T)^n \cdot y \\ \end{bmatrix}$$

This vector is exactly $A^T y$. Therefore we have $(Ax) \cdot y = x \cdot (A^T y)$, as desired.

Problem C: Since $z = Ay$ for some $y \in \mathbf{R}^n$, we have $x \cdot z = z \cdot x = (Ay) \cdot x$, which is equal to $y \cdot (A^T x)$ by the previous exercise. But $A^T x = 0$, so that $x \cdot z = y \cdot 0$, which is zero, since the dot product of any vector with the zero vector yields zero.

Problem D: Suppose $\alpha = 0$. We will show that H is a subspace of $\mathbf{R}^n$. Let $v_1, v_2 \in H$. Then we know $v_1 \cdot x_0 = v_2 \cdot x_0 = 0$. Now let $c_1, c_2 \in \mathbf{R}$, and consider $v = c_1 v_1 + c_2 v_2$. We have $v \cdot x_0 = (c_1 v_1 + c_2 v_2) \cdot x_0 = c_1(v_1 \cdot x_0) + c_2(v_2 \cdot x_0) = c_1 \cdot 0 + c_2 \cdot 0 = 0$, so that $v \in H$. Therefore H is closed under linear combinations, so that H is a subspace of $\mathbf{R}^n$. Now assume that H is a subspace of $\mathbf{R}^n$. Then we must have $0 \in H$, for H must be closed under scalar multiplication, so if we take any $v \in H$, then we must have $cv \in H$ for any $c \in \mathbf{R}$. In particular, if $c = 0$, so that $cv = 0$, then we have $0 \in H$. But if the zero vector belongs to H, then we have $0 \cdot x_0 = \alpha$. But since $0 \cdot x_0 = 0$, we must have $\alpha = 0$.

Problem E:

(a) We showed in class that if Q is an orthogonal matrix, then $Q^T Q = I$. Then using Problem A, $(Qx) \cdot (Qy) = x \cdot (Q^T(Qy)) = x \cdot ((Q^T Q)y) = x \cdot (Iy) = x \cdot y$, as desired.

(b) Using part (a) with $x = y$, we have $\|Qx\| = \sqrt{(Qx) \cdot (Qx)} = \sqrt{x \cdot x} = \|x\|$.

(c) Let x be an eigenvector of Q, associated with the eigenvalue λ. Then $x \neq 0$, and $Qx = \lambda x$. Therefore, $\|Qx\| = \|\lambda x\| = |\lambda|\|x\|$, using part (ii) of Proposition 2.13. But also, $\|Qx\| = \|x\|$ by part (b); hence $\|x\| = |\lambda|\|x\|$. Since $\|x\| \neq 0$, the only way this last equation can hold is if $|\lambda| = 1$, and if λ is real, then we must have $\lambda = 1$ or $\lambda = -1$.