18.303 Problem Set 1 Solutions

Problem 1: (5+(2+2+2+2+2)+(10+5) points)

Note that I don't expect you to rederive basic linear-algebra facts. You can use things derived in 18.06, like the existence of an orthonormal diagonalization of Hermitian matrices.

(a) Since it is Hermitian, B can be diagonalized: $B = Q\Lambda Q^*$, where Q is the matrix whose columns are the eigenvectors (chosen orthonormal so that $Q^{-1} = Q^*$) and Λ is the diagonal matrix of eigenvalues. Define $\sqrt{\Lambda}$ as the diagonal matrix of the (positive) square roots of the eigenvalues, which is possible because the eigenvalues are > 0 (since B is positive-definite). Then define $\sqrt{B} = Q\sqrt{\Lambda}Q^*$, and by inspection we obtain $(\sqrt{B})^2 = B$. By construction, $\sqrt{B}$ is positive-definite and Hermitian.

It is easy to see that this $\sqrt{B}$ is unique, even though the eigenvectors X are not unique, because any acceptable transformation of Q must commute with Λ and hence with $\sqrt{\Lambda}$. Consider for simplicity the case of distinct eigenvalues: in this case, we can only scale the eigenvectors by (nonzero) constants, corresponding to multiplying Q on the right by a diagonal (nonsingular) matrix D. This gives the same B for any D, since $QD\Lambda(QD)^{-1} = Q\Lambda DD^{-1}Q^{-1} = Q\Lambda Q^{-1}$ (diagonal matrices commute), and for the same reason it gives the same $\sqrt{B}$. For repeated eigenvalues λ, D can have off-diagonal elements that mix eigenvectors of the same eigenvalue, but D still commutes with Λ because these off-diagonal elements only appear in blocks where Λ is a multiple λI of the identity (which commutes with anything).

(b) Solutions:

(i) From 18.06, $B^{-1}A$ is similar to $C = MB^{-1}AM^{-1}$ for any invertible M. Let $M = B^{1/2}$ from above. Then $C = B^{-1/2}AB^{-1/2}$, which is clearly Hermitian since A and $B^{-1/2}$ are Hermitian. (Why is $B^{-1/2}$ Hermitian? Because $B^{1/2}$ is Hermitian from above, and the inverse of a Hermitian matrix is Hermitian.)

(ii) From 18.06, similarity means that $B^{-1}A$ has the same eigenvalues as C, and since C is Hermitian these eigenvalues are real.

(iii) No, they are not (in general) orthogonal. The eigenvectors Q of C are (or can be chosen) to be orthonormal ($Q^*Q = I$), but the eigenvectors of $B^{-1}A$ are $X = M^{-1}Q = B^{-1/2}Q$, and hence $X^*X = Q^*B^{-1}Q \neq I$ unless $B = I$.

(iv) Note that there was a typo in the pset. The `eigvals` function returns only the eigenvalues; you should use the `eig` function instead to get both eigenvalues and eigenvectors, as explained in the Julia handout.

The array `lambda` that you obtain in Julia should be purely real, as expected. (You might notice that the eigenvalues are in somewhat random order, e.g. I got -8.11,3.73,1.65,-1.502,0.443. This is a side effect of how eigenvalues of non-symmetric matrices are computed in standard linear-algebra libraries like LAPACK.) You can check orthogonality by computing X^*X via `X'*X`, and the result is not a diagonal matrix (or even close to one), hence the vectors are not orthogonal.

(v) When you compute $C = X^*BX$ via `C=X'*B*X`, you should find that C is nearly diagonal: the off-diagonal entries are all very close to zero (around 10^{-15} or less). They would be *exactly* zero except for roundoff errors (as mentioned in class, computers keep only around 15 significant digits). From the definition of matrix multiplication, the entry C_{ij} is given by the i-th row of X^* multiplied by B, multiplied by the j-th column of X. But the j-th column X is the j-th eigenvector $\mathbf{x}_j$, and the i-th row of X^* is $\mathbf{x}_i^*$. Hence $C_{ij} = \mathbf{x}_i^* B\mathbf{x_j}$, which looks like a dot product but with B in the middle. The fact that C

is diagonal means that $\mathbf{x}_i^* B \mathbf{x_j} = 0$ for $i \neq j$, which is a kind of orthogonality relation.

[In fact, if we define the inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* B \mathbf{y}$, this is a perfectly good inner product (it satisifies all the inner-product criteria because B is positive-definite), and we will see in the next pset that $B^{-1}A$ is actually self-adjoint under this inner product. Hence it is no surprise that we get real eigenvalues and orthogonal eigenvectors with respect to this inner product.]

(c) Solutions:

(i) If we write $\mathbf{x}(t) = \sum_{n=1}^4 c_n(t)\mathbf{x}_n$, then plugging it into the ODE and using the eigenvalue equation yields

$$\sum_{n=1}^{4} [\ddot{c}_n - 2\dot{c}_n - \lambda_n c]\, \mathbf{x}_n = 0.$$

Using the fact that the $\mathbf{x}_n$ are necessarily **orthogonal** (they are eigenvectors of a Hermitian matrix for distinct eigenvalues), we can take the dot product of both sides with $\mathbf{x}_m$ to find that $\ddot{c}_n - 2\dot{c}_n - \lambda_n c = 0$ for each n, and hence

$$c_n(t) = \alpha_n e^{(1+\sqrt{1+\lambda_n})t} + \beta_n e^{(1-\sqrt{1+\lambda_n})t}$$

for constants α_n and β_n to be determined from the initial conditions. Plugging in the initial conditions $\mathbf{x}(0) = \mathbf{a}_0$ and $\mathbf{x}'(0) = \mathbf{b}_0$, we obtain the equations:

$$\sum_{n=1}^{4} (\alpha_n + \beta_n)\mathbf{x}_n = \mathbf{a}_0,$$

$$\sum_{n=1}^{4} ([\alpha_n + \beta_n] + \sqrt{1+\lambda_n}[\alpha_n - \beta_n])\mathbf{x}_n = \mathbf{b}_0.$$

Again using orthogonality to pull out the n-th term, we find

$$\alpha_n + \beta_n = \frac{\mathbf{x}_n^* \mathbf{a}_0}{\|\mathbf{x}_n\|^2}$$

$$[\alpha_n + \beta_n] + \sqrt{1+\lambda_n}[\alpha_n - \beta_n] = \frac{\mathbf{x}_n^* \mathbf{b}_0}{\|\mathbf{x}_n\|^2} \implies \alpha_n - \beta_n = \frac{\mathbf{x}_n^*(\mathbf{b}_0 - \mathbf{a}_0)}{\|\mathbf{x}_n\|^2 \sqrt{1+\lambda_n}}$$

(note that we were *not* given that $\mathbf{x}_n$ were normalized to unit length, and this is *not* automatic) and hence we can solve for α_n and β_n to obtain:

$$\mathbf{x}(t) = \sum_{n=1}^{4} \left(\left[\mathbf{x}_n^* \mathbf{a}_0 + \frac{\mathbf{x}_n^*(\mathbf{b}_0 - \mathbf{a}_0)}{\sqrt{1+\lambda_n}} \right] e^{(1+\sqrt{1+\lambda_n})t} + \left[\mathbf{x}_n^* \mathbf{a}_0 + \frac{\mathbf{x}_n^*(\mathbf{b}_0 - \mathbf{a}_0)}{\sqrt{1+\lambda_n}} \right] e^{(1-\sqrt{1+\lambda_n})t} \right) \frac{\mathbf{x}_n}{2\|\mathbf{x}_n\|^2}.$$

(ii) After a long time, this expression will be dominated by the fastest growing term, which is the $e^{(1+\sqrt{1+\lambda_n})t}$ term for $\lambda_4 = 24$, hence:

$$\mathbf{x}(t) \approx \left[\mathbf{x}_4^* \mathbf{a}_0 + \frac{\mathbf{x}_4^*(\mathbf{b}_0 - \mathbf{a}_0)}{5} \right] e^{6t} \frac{\mathbf{x}_4}{2\|\mathbf{x}_4\|^2}.$$

Problem 2: ((5+5+10)+5+5 points)

(a) Suppose that we we change the boundary conditions to the *periodic* boundary condition $u(0) = u(L)$.

(i) As in class, the eigenfunctions are sines, cosines, and exponentials, and it only remains to apply the boundary conditions. $\sin(kx)$ is periodic if $k = \frac{2\pi n}{L}$ for $n = 1, 2, \ldots$ (excluding $n = 0$ because we do not allow zero eigenfunctions and excluding $n < 0$ because they are not linearly independent), and $\cos(kx)$ is periodic if $n = 0, 1, 2, \ldots$ (excluding $n < 0$ since they are the same functions). The eigenvalues are $-k^2 = -(2\pi n/L)^2$.

e^{kx} is periodic only for imaginary $k = i\frac{2\pi n}{L}$, but in this case we obtain $e^{i\frac{2\pi n}{L}x} = \cos(2\pi nx/L) + i\sin(2\pi nx/L)$, which is *not linearly independent* of the sin and cos eigenfunctions above. Recall from 18.06 that the eigenvectors for a given eigenvalue form a vector space (the null space of $A - \lambda I$), and when asked for eigenvectors we only want a *basis* of this vector space. Alternatively, it is acceptable to start with exponentials and call our eigenfunctions $e^{i\frac{2\pi n}{L}x}$ for all integers n, in which case we wouldn't give sin and cos eigenfunctions separately.

Similarly, $\sin(\phi + 2\pi nx/L)$ is periodic for any ϕ, but this is not linearly independent since $\sin(\phi + 2\pi nx/L) = \sin\phi\cos(2\pi nx/L) + \cos\phi\sin(2\pi nx/L)$.

[Several of you were tempted to also allow $\sin(m\pi x/L)$ for **odd** m (not just the even m considered above). At first glance, this seems like it satisfies the PDE and also has $u(0) = u(L)$ $(= 0)$. Consider, for example, $m = 1$, i.e. $\sin(\pi x/L)$ solutions. This can't be right, however; e.g. it is not orthogonal to $1 = \cos(0x)$, as required for self-adjoint problems. The basic problem here is that if you consider the periodic extension of $\sin(\pi x/L)$, then it doesn't actually satisfy the PDE, because it has a slope discontinuity at the endpoints. Another way of thinking about it is that periodic boundary conditions arise because we have a PDE defined on a torus, e.g. diffusion around a circular tube, and in this case the choice of endpoints is not unique—we can easily redefine our endpoints so that $x = 0$ is in the "middle" of the domain, making it clearer that we can't have a kink there. (This is one of those cases where to be completely rigorous we would need to be a bit more careful about defining the domain of our operator.)]

(ii) No, any solution will not be unique, because we now have a nonzero nullspace spanned by the constant function $u(x) = 1$ (which is periodic): $\frac{d^2}{dx^2}1 = 0$. Equivalently, we have a 0 eigenvalue corresponding to $\cos(2\pi nx/L)$ for $n = 0$ above.

(iii) As suggested, let us restrict ourselves to $f(x)$ with a convergent Fourier series. That is, as in class, we are expanding $f(x)$ in terms of the eigenfunctions:

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{i\frac{2\pi n}{L}x}.$$

(You could also write out the Fourier series in terms of sines and cosines, but the complex-exponential form is more compact so I will use it here.) Here, the coefficients c_n, by the usual orthogonality properties of the Fourier series, or equivalently by self-adjointness of $\hat{A}$, are $c_n = \frac{1}{L}\int_0^L e^{-\frac{2\pi n}{L}x} f(x)dx$.

In order to solve $\frac{d^2u}{dx^2} = f$, as in class we would divide each term by its eigenvalue $-(2\pi n/L)^2$, but we can only do this for $n \neq 0$. Hence, we can only solve the equation if the $n = 0$ term is absent, i.e. $c_0 = 0$. Appling the explicit formula for c_0, the equation is solvable (for f with a Fourier series) if and only if:

$$\boxed{\int_0^L f(x)dx = 0}.$$

There are other ways to come to the same conclusion. For example, we could expand $u(x)$ in a Fourier series (i.e. in the eigenfunction basis), apply d^2/dx^2, and ask *what is the column space* of d^2/dx^2? Again, we would find that upon taking the second derivative the $n = 0$ (constant) term vanishes, and so the column space consist of Fourier series missing a constant term.

The same reasoning works if you write out the Fourier series in terms of sin and cos sums separately, in which case you find that f must be missing the $n = 0$ cosine term, giving the same result.

(b) No. For example, the function 0 (which must be in any vector space) does not satisy those boundary conditions. (Also adding functions doesn't work, scaling them by constants, etcetera.)

(c) We merely pick any twice-differentiable function $q(x)$ with $q(L) - q(0) = -1$, in which case $u(L) - u(0) = [v(L) - v(0)] + [q(L) - q(0)] = 1 - 1 = 0$ and u is periodic. Then, plugging $v = u - q$ into $\frac{d^2}{dx^2}v(x) = f(x)$, we obtain

$$\frac{d^2}{dx^2}u(x) = f(x) + \frac{d^2q}{dx^2},$$

which is the (periodic-u) Poisson equation for u with a (possibly) modified right-hand side.

For example, the simplest such q is probably $q(x) = x/L$, in which case $d^2q/dx^2 = 0$ and u solves the Poisson equation with an *unmodified* right-hand side.

Problem 3: (10+10 points)

We are using a difference approximation of the form:

$$u'(x) \approx \frac{-u(x + 2\Delta x) + c \cdot u(x + \Delta x) - c \cdot u(x - \Delta x) + u(x - 2\Delta x)}{d \cdot \Delta x}.$$

(a) First, we Taylor expand:

$$u(x + \Delta x) = \sum_{n=0}^{\infty} \frac{u^{(n)}(x)}{n!}\Delta x^n.$$

The numerator of the difference formula flips sign if $\Delta x \to -\Delta x$, which means that when you plug in the Taylor series all of the even powers of Δx must cancel! To get 4th-order accuracy, the Δx^3 term in the numerator (which would give an error $\sim \Delta x^2$) must cancel as well, and this determines our choice of c: the Δx^3 term in the numerator is

$$\frac{u'''(x)}{3!}\Delta x^3 \left[-2^3 + c + c - 2^3\right],$$

and hence we must have $\boxed{c = 2^3 = 8}$. The remaining terms in the numerator are the Δx term and the Δx^5 term:

$$u'(x)\Delta x\left[-2 + c + c - 2\right] + \frac{u^{(5)}(x)}{5!}\Delta x^5\left[-2^5 + c + c - 2^5\right] = 12u'(x)\Delta x - \frac{2}{5}u^{(5)}(x)\Delta x^5 + \cdots.$$

Clearly, to get the correct $u'(x)$ as $\Delta x \to 0$, we must have $\boxed{d = 12}$. Hence, the error is approximately $-\frac{1}{30}u^{(5)}(x)\Delta x^4$, which is $\sim \Delta x^4$ as desired.

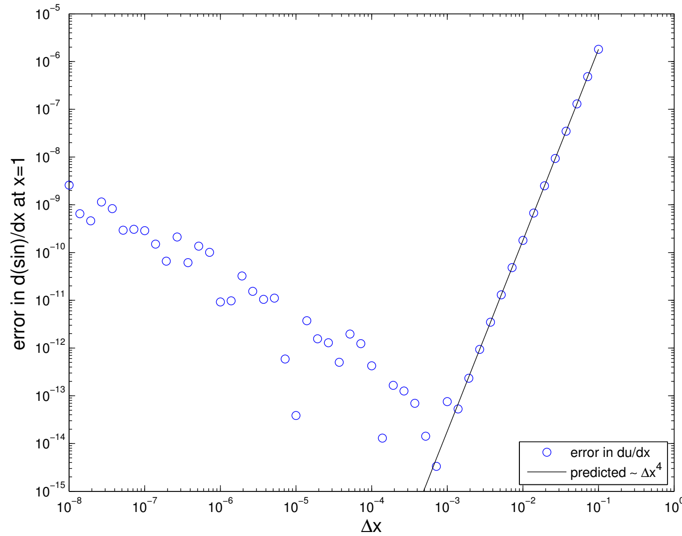

Figure 1: Actual vs. predicted error for problem 1(b), using fourth-order difference approximation for $u'(x)$ with $u(x) = \sin(x)$, at $x = 1$.

(b) The Julia code is the same as in the handout, except now we compute our difference approximation by the command: d = `(-sin(x+2*dx) + 8*sin(x+dx) - 8*sin(x+dx) + sin(x-2*dx)) ./ (12 * dx)`; the result is plotted in Fig. 1. Note that the error falls as a straight line (a power law), until it reaches $\sim 10^{-15}$, when it starts becoming dominated by roundoff errors (and actually gets worse). To verify the order of accuracy, it would be sufficient to check the slope of the straight-line region, but it is more fun to plot the actual predicted error from the previous part, where $\frac{d^5}{dx^5}\sin(x) = -\cos(x)$. Clearly the predicted error is almost exactly right (until roundoff errors take over).