

## 18.335 Problem Set 4 Solutions

### Problem 1: ((5+5)+5+5 points)

Consider the “generalized” eigenvalue problem  $Ax = \lambda Bx$  where  $A$  and  $B$  are Hermitian and  $B$  is positive-definite. You could turn this into an ordinary eigenvalue problem  $B^{-1}Ax = \lambda x$ , but then you disguise the Hermitian nature of the underlying problem.

- (a) If  $Ax = \lambda Bx$  for  $x \neq 0$ , then  $x^*Ax = \lambda x^*Bx = (Ax)^*x = (\lambda Bx)^*x = \bar{\lambda} x^*Bx$ , and since  $x^*Bx > 0$  (because  $B$  is positive-definite and  $x \neq 0$ ), it follows that  $\lambda = \bar{\lambda}$  (**real**).

Suppose  $Ax_1 = \lambda_1 Bx_1$  and  $Ax_2 = \lambda_2 Bx_2$  with  $\lambda_1 \neq \lambda_2$ . Then  $x_1^*Ax_2 = \lambda_2 x_1^*Bx_2 = (Ax_1)^*x_2 = (\lambda_1 Bx_1)^*x_2 = \lambda_1 x_1^*Bx_2$ , hence  $(\lambda_2 - \lambda_1)x_1^*Bx_2 = 0$ , and hence  $x_1^*Bx_2 = 0$  (since  $\lambda_2 - \lambda_1 \neq 0$ ).

(Even if  $\lambda_1 = \lambda_2$  and  $x_1$  and  $x_2$  are linearly independent, we can choose an orthonormal basis for the span of  $x_1$  and  $x_2$  — all linear combinations are also eigenvectors with the same  $\lambda$  — by e.g. Gram–Schmidt. Hence we can choose the eigenvectors to be orthonormal under the  $B$ -weighted inner product.)

- (b) From the previous part, we can choose the eigenvectors  $X$  to be  $B$ -orthonormal in the sense that  $X^*BX = I$ , and hence  $X^{-1} = X^*B$ . Since  $Ax = \lambda Bx \implies AX = BX\Lambda$ , it follows that  $A = BX\Lambda X^{-1} = \boxed{BX\Lambda X^*B}$ .
- (c) Since  $B$  is Hermitian positive-definite, it has a Cholesky factorization  $B = R^*R$  (where  $R$  is upper-triangular). Therefore, take  $Ax = \lambda Bx = \lambda R^*(Rx)$  and let  $\boxed{Rx = y}$ . It follows that  $AR^{-1}y = \lambda R^*y$ , or equivalently  $Hy = \lambda y$  where  $\boxed{H = (R^*)^{-1}AR^{-1} = (R^{-1})^*AR^{-1}}$  is Hermitian. (A common notation for  $(R^*)^{-1} = (R^{-1})^*$  is  $R^{-*}$ .)

Note that this is quite a practical algorithm, since Cholesky factorization is cheap compared to finding eigenvalues, and forming  $(R^{-1})^*AR^{-1}$  can be done using a sequence of back/forward-substitutions since  $R$  is upper-triangular. And given an eigenvector  $y$  of  $H$  we can find the corresponding generalized eigenvector  $x$  by a back-substitution solve.

### Problem 2 (15 points)

Trefethen, problem 27.5. The basic answer here is that there is a big roundoff error, but it is in the direction of the eigenvector we want, so we don't care (since overall scale factors are irrelevant to eigenvector computations).

In finite precision, instead of  $w = A^{-1}v$ , we will get  $\tilde{w} = w + \delta w$  where  $\delta w = -(A + \delta A)^{-1}\delta A w$  (from the formula on page 95), where  $\delta A = O(\epsilon_{\text{machine}})\|A\|$  is the backwards error. [Note that we cannot use  $\delta w \approx -A^{-1}\delta A w$ , which neglects the  $\delta A \delta w$  terms, because in this case  $\delta w$  is not small.] The key point, however, is to show that  $\delta w$  is mostly parallel to  $q_1$ , the eigenvector corresponding to the smallest-magnitude eigenvalue  $\lambda_1$  (it is given that all other eigenvalues have magnitude  $\geq |\lambda_2| \gg |\lambda_1|$ ). Since  $w$  is also mostly parallel to  $q_1$ , this will mean that  $\tilde{w}/\|\tilde{w}\|_2 \approx q_1 \approx w/\|w\|_2$ .

First, exactly as in our analysis of the power method, note that  $w = A^{-1}v = \alpha_1 q_1 [1 + O(\lambda_1/\lambda_2)]$ , since  $A^{-1}$  amplifies the  $q_1$  component of  $v$  by  $1/|\lambda_1|$  which is much bigger than the inverse of all the other eigenvalues. Thus,  $w/\|w\|_2 = q_1 [\pm 1 + O(\lambda_1/\lambda_2)]$ , where the  $\pm 1$  is an arbitrary phase (or  $e^{i\phi}$  if we are talking about complex vectors).

Second, if we Taylor-expand  $(A + \delta A)^{-1}$  in powers of  $\delta A$ , i.e. in powers of  $\epsilon_{\text{machine}}$ , we obtain:<sup>1</sup>  $(A + \delta A)^{-1} = A^{-1} - A^{-1}\delta A A^{-1} + O(\epsilon_{\text{machine}}^2)$ . Since all of the terms in this expansion are multiplied

<sup>1</sup>Write  $(A + \delta A)^{-1} = [A(I + A^{-1}\delta A)]^{-1} = (I + A^{-1}\delta A)^{-1}A^{-1} \approx (I - A^{-1}\delta A)A^{-1} = A^{-1} - A^{-1}\delta A A^{-1}$ . Another approach is to let  $B = (A + \delta A)^{-1} = B_0 + B_1 + \dots$  where  $B_k$  is the  $k$ -th order term in  $\delta A$ , collect terms order-by-order in  $I = (B_0 + B_1 + \dots)(A + \delta A) = B_0A + (B_0\delta A + B_1A) + \dots$ , and you immediately find that  $B_0 = A^{-1}$ ,  $B_1 = -B_0\delta A A^{-1} = -A^{-1}\delta A A^{-1}$ , and so on.

on the *left* by  $A^{-1}$ , when multiplied by *any* vector they will again amplify the  $q_1$  component much more than any other component. In particular, the vector  $\delta A w$  is a vector in a random direction (since  $\delta A$  comes from roundoff and is essentially random) and hence will have some nonzero  $q_1$  component. Thus,  $\delta w = -(A + \delta A)^{-1} \delta A w = \beta_1 q_1 [1 + O(\lambda_1/\lambda_2)]$  for some constant  $\beta_1$ .

Putting these things together, we see that  $\tilde{w} = (\alpha_1 + \beta_1) q_1 [1 + O(\lambda_1/\lambda_2)]$ , and hence  $\tilde{w}/\|\tilde{w}\|_2 = q_1 [\pm 1 + O(\lambda_1/\lambda_2)] = \frac{w}{\|w\|_2} [\pm 1 + O(\lambda_1/\lambda_2)]$ . Q.E.D.

### Problem 3 (5+5+5+5+5 pts):

Trefethen, problem 33.2:

- (a) In this case, the  $q_{n+1}$  vector is multiplied by a zero row in  $\tilde{H}_n$ , and we can simplify 33.13 to  $AQ_n = Q_n H_n$ . If we consider the full Hessenberg reduction,  $H = Q^* A Q$ , it must have a “block Schur” form:

$$H = \begin{pmatrix} H_n & B \\ 0 & H' \end{pmatrix},$$

where  $H'$  is an  $(m-n) \times (m-n)$  upper-Hessenberg matrix and  $B \in \mathbb{C}^{n \times (m-n)}$ . (It is *not* necessarily the case that  $B = 0$ ; this is only true if  $A$  is Hermitian.)

- (b)  $Q_n$  is a basis for  $\mathcal{K}_n$ , so any vector  $x \in \mathcal{K}_n$  can be written as  $x = Q_n y$  for some  $y \in \mathbb{C}^n$ . Hence, from above,  $Ax = AQ_n y = Q_n H_n y = Q_n (H_n y) \in \mathcal{K}_n$ . Q.E.D.
- (c) The  $(n+1)$  basis vector,  $A^n b$ , is equal to  $A(A^{n-1} b)$  where  $A^{n-1} b \in \mathcal{K}_n$ . Hence, from above,  $A^n b \in \mathcal{K}_n$  and thus  $\mathcal{K}_{n+1} = \mathcal{K}_n$ . By induction,  $\mathcal{K}_\ell = \mathcal{K}_n$  for  $\ell \geq n$ .
- (d) If  $H_n y = \lambda y$ , then  $AQ_n y = Q_n H_n y = \lambda Q_n y$ , and hence  $\lambda$  is an eigenvalue of  $A$  with eigenvector  $Q_n y$ .
- (e) If  $A$  is nonsingular, then  $H_n$  is nonsingular (if it had a zero eigenvalue,  $A$  would too from above). Hence, noting that  $b$  is proportional to the first column of  $Q_n$ , we have:

$$x = A^{-1} b = A^{-1} Q_n e_1 \|b\| = A^{-1} Q_n H_n H_n^{-1} e_1 \|b\| = A^{-1} A Q_n H_n^{-1} e_1 \|b\| = Q_n H_n^{-1} e_1 \|b\| \in \mathcal{K}_n.$$

Q.E.D.