

4.6.1 Prove that if  $A$  is diagonally dominant and if  $Q$  is chosen as in the Jacobi method, then

$$\rho(I - Q^{-1}A) < 1.$$

*Proof.* If  $Q$  simply consists of the diagonal entries of  $A$  then  $Q^{-1}$  acts on  $a_{i,j}$  by  $a_{i,j} \mapsto a_{i,j}/a_{i,i}$ . It follows that the diagonal entries of  $Q^{-1}A$  are all 1 and, since the ratios between entries in the same row remain unchanged, by  $A$ 's diagonal dominance

$$(Q^{-1}A)_{i,i} = 1 = \frac{|a_{i,i}|}{|a_{i,i}|} > \left( \sum_{j \neq i} |a_{i,j}| \right) / |a_{i,i}| = \sum_{j \neq i} \frac{|a_{i,j}|}{|a_{i,i}|} = \sum_{j \neq i} |(Q^{-1}A)_{i,j}|$$

and so  $Q^{-1}A$  is also diagonally dominant. It follows that  $\|I - Q^{-1}A\|_{\infty} < 1$ , and  $\rho(I - Q^{-1}A) \leq \|I - Q^{-1}A\|_{\infty}$  by definition of infimum.  $\square$

4.6.5 Let  $\|\cdot\|$  be a norm on  $\mathbb{R}^n$ . Let  $S$  be an  $n \times n$  nonsingular matrix. Define  $\|x\|' = \|Sx\|$ . Prove that  $\|\cdot\|'$  is a norm.

*Proof.* Non-degeneracy is clear as  $x \neq 0 \Rightarrow Sx \neq 0 \Rightarrow \|Sx\| > 0$  and  $\|Sx\| = 0 \iff Sx = 0 \iff x = 0$ .

Absolute homogeneity:

$$\|\lambda x\|' = \|S(\lambda x)\| = \|\lambda(Sx)\| = |\lambda| \|Sx\| = |\lambda| \|x\|'.$$

For triangle inequality:

$$\|x + y\|' = \|S(x + y)\| = \|Sx + Sy\| \leq \|Sx\| + \|Sy\| = \|x\|' + \|y\|'. \quad \square$$

4.6.6 Let  $\|\cdot\|$  be a subordinate matrix norm and let  $S$  be a nonsingular matrix. Define  $\|A\|' = \|SAS^{-1}\|$ . Show that  $\|\cdot\|'$  is a subordinate matrix norm.

*Proof.*  $\|A\|' = \|SAS^{-1}\| = \sup_{\|x\|=1} \|SAS^{-1}x\| = \sup_{\|Sx\|=1} \|SAx\| = \sup_{\|x\|=1} \|Ax\|'$ , where the second step takes advantage of the fact  $SS^{-1}x = x$  and the third uses the result from 4.6.5.  $\square$

tage of the fact  $SS^{-1}x = x$  and the third uses the result from 4.6.5.  $\square$

4.6.10 Which of the norm axioms are satisfied by the spectral radius function  $\rho$  and which are not?

### Solution

Non-degeneracy is not satisfied: the eigenvalues of the nonzero matrix  $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$  are both 0 so  $\rho(\cdot) = 0$ .

Absolute homogeneity is not satisfied: suppose  $\lambda_i$  is the maximal eigenvalue of  $A$  with  $Ax_i = \lambda_i x_i$ . We then have

$$(-A)x_i = -(Ax_i) = -\lambda_i x_i \neq |-1| \cdot \lambda_i x_i.$$

Triangle inequality is not satisfied, either. Consider  $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$  and  $B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ . Clearly  $\rho(A) =$

$\rho(B) = 0$  but  $A + B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$  which has  $\rho(\cdot) = 1$ .

4.6.16 Prove that if  $A$  is nonsingular then  $AA^*$  is positive definite.

*Proof.*  $A^*$  is also nonsingular. Thus, for  $x \neq 0$ ,  $A^*x \neq 0$  and  $x^*AA^*x = (A^*x)^*(A^*x) = \sqrt{\|A^*x\|} > 0$ .  $\square$

4.6.18 Prove that if  $A$  is positive definite, then so are  $A^2, A^3, \dots$  as well as  $A^{-1}, A^{-2}, \dots$

*Proof.* If  $A$  is positive definite then its eigenvalues are strictly positive. Notice that if  $Ax = \lambda x$  then  $A^k x = A^{k-1}(Ax) = A^{k-1}\lambda x = \dots = \lambda^k x$ . Thus all eigenvalues of  $A^k$  are also positive, and this implies the positive definiteness of  $A^k$ : for an eigenvector  $x$  corresponding to eigenvalue  $\lambda > 0$  of  $B$ , we have

$$x^* Bx = x^*(\lambda x) = \lambda(x^* x) > 0,$$

and the result for more general  $x$ 's (i.e., linear combination of eigenvectors) follow from linearity of matrix operation. For powers of  $A^{-1}$ , simply notice that the eigenvalues of  $A^{-1}$  are  $1/\lambda_i$ 's.  $\square$

5.1.1 Let  $A$  be an  $n \times n$  matrix with  $P$ 's column being its linearly independent eigenvectors. What is  $P^{-1}AP$ ?

**Solution**

$$P^{-1}AP = P^{-1}(P\Lambda P^{-1})P = \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n).$$

5.1.3 In the power method, let  $r_k = \varphi(x^{(k+1)})/\varphi(x^{(k)})$ . We know that  $\lim_{k \rightarrow \infty} r_k = \lambda_1$ . Show that the relative errors obey

$$\frac{r_k - \lambda_1}{\lambda_1} = \left(\frac{\lambda_2}{\lambda_1}\right)^k c_k$$

where  $\{c_k\}$  is convergent.

*Proof.* Since

$$r_k = \frac{\varphi(x^{(k+1)})}{\varphi(x^{(k)})} = \frac{\lambda_1^{k+1}(\varphi(u^{(1)}) + \varphi(\epsilon^{(k+1)}))}{\lambda_1^k(\varphi(u^{(1)}) + \varphi(\epsilon^{(k)}))} = \lambda_1 \frac{\varphi(u^{(1)}) + \varphi(\epsilon^{(k+1)})}{\varphi(u^{(1)}) + \varphi(\epsilon^{(k)})},$$

we have

$$\begin{aligned} \frac{r_k - \lambda_1}{\lambda_1} &= \frac{\varphi(u^{(1)}) + \varphi(\epsilon^{(k+1)})}{\varphi(u^{(1)}) + \varphi(\epsilon^{(k)})} - 1 = \frac{\varphi(\epsilon^{(k+1)}) - \varphi(\epsilon^{(k)})}{\varphi(u^{(1)}) + \varphi(\epsilon^{(k)})} \\ &= \frac{\varphi[(\lambda_2/\lambda_1)^{k+1}u^{(2)} - (\lambda_2/\lambda_1)^k u^{(2)} + \dots + (\lambda_n/\lambda_1)^{k+1}u^{(n)} - (\lambda_n/\lambda_1)^k u^{(n)}]}{\varphi(u^{(1)} + \epsilon^{(k)})} \\ &= \frac{\varphi[(\lambda_2/\lambda_1)^k(\lambda_2/\lambda_1 - 1)u^{(2)} + \dots]}{\varphi(u^{(1)} + \epsilon^{(k)})} \\ &= \left(\frac{\lambda_2}{\lambda_1}\right)^k \frac{\varphi[(\lambda_2/\lambda_1 - 1)u^{(2)} + (\lambda_3/\lambda_2)^k(\lambda_3/\lambda_1 - 1)u^{(3)} + \dots]}{\varphi(u^{(1)} + \epsilon^{(k)})} \\ &= \left(\frac{\lambda_2}{\lambda_1}\right)^k \frac{\varphi(\lambda_2/\lambda_1 - 1)u^{(2)}}{\varphi(u^{(1)})}. \end{aligned} \quad \square$$

5.1.4 Show that  $r_{k+1} - \lambda = (c + \delta_k)(r_k - \lambda_1)$  where  $|c| < 1$  and  $\lim_{n \rightarrow \infty} \delta_k = 0$  so that the Aitken acceleration is applicable.

*Proof.* First we compute  $r_{k+1} - \lambda_1$ :

$$\begin{aligned} r_{k+1} - \lambda_1 &= \frac{\varphi(x^{(k+2)}) - \lambda_1 \varphi(x^{(k+1)})}{\varphi(x^{(k+1)})} = \frac{\varphi(x^{(k+2)}) - \lambda_1 x^{(k+1)}}{\varphi(x^{(k+1)})} \\ &= \frac{\varphi[\lambda_1^{k+1}(a_1 u^{(1)} + \epsilon^{(k+2)}) - \lambda_1^{k+2}(a_1 u^{(1)} + \epsilon^{(k+1)})]}{\varphi(\lambda_1^{k+1}(a_1 u^{(1)} + \epsilon^{(k+1)}))} \\ &= \frac{\lambda_1^{k+2} \varphi(\epsilon^{(k+2)} - \epsilon^{(k+1)})}{\lambda_1^{k+1} \varphi(a_1 u^{(1)} + \epsilon^{(k+1)})} \\ &= \lambda_1 \frac{\varphi(\epsilon^{(k+2)} - \epsilon^{(k+1)})}{\varphi(a_1 u^{(1)} + \epsilon^{(k+1)})}. \end{aligned}$$

Therefore

$$\frac{r_{k+1} - \lambda_1}{r_k - \lambda_1} = \frac{\varphi(a_1 u^{(1)} + \epsilon^{(k)})}{\varphi(\epsilon^{(k+1)} - \epsilon^{(k)})} \cdot \frac{\varphi(\epsilon^{(k+2)} - \epsilon^{(k-1)})}{\varphi(a_1 u^{(1)} + \epsilon^{(k+1)})}$$

which indeed converges to a constant as a constant with absolute value  $< 1$  as  $k \rightarrow \infty$ .  $\square$

5.1.9 What can you prove about Aitken acceleration if the sequence  $\{r_n\}$  satisfies only the hypothesis  $|r_{n+1} - r| \leq c|r_n - r|$  with  $0 < c < 0.2$ ?

**Solution**

The faster convergence by Aitken acceleration still holds as certainly any sequence bounded by  $(-0.2, 0.2)$  can be written as  $a_n = c' + \delta_n$  for some fixed  $c'$  and  $\delta_n \rightarrow 0$ .