GGSSS13 Solutions to Problems Higham

1. Consider two Jordan canonical forms

$$A = ZJZ^{-1} = WJW^{-1} \tag{1}$$

(by incorporating a permutation matrix in W we can assume without loss of generality that J is the same matrix in both cases). The definition gives $f_1(A) = Zf(J)Z^{-1}$, $f_2(A) = Wf(J)W^{-1}$ and we need to show that $f_1(A) = f_2(A)$, that is, $W^{-1}Zf(J)Z^{-1}W = f(J)$, or $X^{-1}f(J)X = f(J)$ where $X = Z^{-1}W$. Now by (1) we have $X^{-1}JX = J$, which implies $f(J) = f(X^{-1}JX) = X^{-1}f(J)X$, the last equality following from the JCF definition. Hence $f_1(A) = f_2(A)$, as required.

2.

(a) The Jordan canonical form of A can be ordered so that $A = ZJZ^{-1}$ with $j_{11} = \lambda$ and $Z(:,1) = x$. Now $f(A) = Zf(J)Z^{-1}$ and so $f(A)Z = Zf(J)$. The first column of this equation gives $f(A)x = f(\lambda)x$, as required. For an alternative proof, let p interpolate to the values of f on the spectrum of A. Then $p(\lambda) = f(\lambda)$. Since $A^k x = \lambda^k x$ for all k, $f(A)x = p(A)x = p(\lambda)x = f(\lambda)x$, as required.

(b) Setting $(\lambda, x) \equiv (\alpha, e)$ in (a) gives the row sum result. If A has column sums α then $A^T e = \alpha e$ and applying the row sum result to A^T gives $f(\alpha)e = f(A^T)e = f(A)^T e$. So $f(A)$ has column sums $f(\alpha)$.

3. It is easiest to use the polynomial interpolation definition, which says that $\cos(\pi A) = p(A)$, where $p(1) = \cos\pi = -1$, $p'(1) = -\pi\sin\pi = 0$, $p(2) = \cos 2\pi = 1$. Writing $p(t) = a + bt + ct^2$ we have

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix},$$

which can be solved to give $p(t) = 1 - 4t + 2t^2$. Hence

$$\cos(\pi A) = p(A) = I - 4A + 2A^2 = \begin{bmatrix} -3 & 0 & 0 & 4 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -2 & 0 & 0 & 3 \end{bmatrix}.$$

Evaluating $\cos(\pi A)$ from its power series would be much more complicated.

4.

For the interpolation definition the result is immediate: $f(A)$ is a polynomial in A and so commutes with B since A does.

For the Cauchy integral definition we have, using $(zI - A)B = B(zI - A)$,

$$\begin{aligned} f(A)B &= \frac{1}{2\pi i}\int_\Gamma f(z)(zI - A)^{-1}B\,dz \\ &= \frac{1}{2\pi i}B\int_\Gamma f(z)(zI - A)^{-1}\,dz = Bf(A). \end{aligned}$$

5. (a) Straightforward. For last part:

$$\begin{aligned} I &= I + AB - (I + AB)A(I + BA)^{-1}B \\ &= I + AB - A(I + BA)(I + BA)^{-1}B \\ &= I + AB - AB \\ &= I \quad \checkmark \end{aligned}$$

(b)

$$\begin{aligned} & (AB)A = A(BA) \\ \Rightarrow\quad & (AB)^2A = ABA(BA) = A(BA)^2. \end{aligned}$$

In general, for any poly p,

$$p(AB)A = Ap(BA).$$

(c) There is a single polynomial p such that $f(AB) = p(AB)$ and $f(BA) = p(BA)$. Hence

$$Af(BA) = Ap(BA) = p(AB)A = f(AB)A.$$

6. Note first that the given assumption on f implies that f is defined on the spectrum of $\alpha I_n + BA$ and at α.

Let $g(t) = f[\alpha + t, \alpha] = t^{-1}(f(\alpha + t) - f(\alpha))$, so that $f(\alpha + t) = f(\alpha) + tg(t)$. Then

$$\begin{aligned} f(\alpha I_m + AB) &= f(\alpha)I_m + ABg(AB) \\ &= f(\alpha)I_m + Ag(BA)B \\ &= f(\alpha)I_m + A(BA)^{-1}(f(\alpha I_n + BA) - f(\alpha)I_n)B, \end{aligned}$$

where the second equality is from the fact that $Bg(AB) = g(BA)B$.

The above formula does not generalize in the proposed way because $h(X) = f(D+X)$ is not a function of X according to our definition—more precisely, it does not correspond to a scalar "stem function" evaluated at X, because of the presence of D. In the "$\alpha I \to D$" generalization, the right-hand side of the formula would contain $f(D_n + BA)$ where D_n is an $n \times n$ diagonal matrix obtained from D, yet there is no reasonable way to define D_n.

7. Define

$$L = \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & -1 & \ddots & \\ & & \ddots & 1 \\ & & & -1 \end{bmatrix} \in \mathbb{R}^{(n+1)\times n}.$$

Then $T_n = L^TL$, $\widetilde{T}_{n+1} = LL^T$.

So $\Lambda(\widetilde{T}_{n+1}) = \Lambda(T_n) \cup \{0\}$ (Strang, 2005).

Example:

```
n = 6; L = gallery('triw',n,-1,1)';
L = L(:,1:n-1), A = L*L', B = L'*L
```

8. Yes for $n \le 2$; no for $n > 2$. AB must have a Jordan form with eigenvalues all zero, these eigenvalues appearing in 1×1 or 2×2 blocks. BA has the same eigenvalues as AB, so the question is whether BA has Jordan blocks of dimension only 1 or 2. This can be answered using Flanders' theorem on the Jordan structures of AB and BA. But working instead from first principles, note that the dimensions of the Jordan blocks cannot exceed 3, because $BABABA = B(ABAB)A = 0$. There are obviously no counterexamples with $n = 2$, but for $n = 3$ we find in MATLAB

```
A =
     0     0     1
     0     0     0
     0     1     0
B =
     0     0     1
     1     0     0
     0     0     0
>> [A*B  A*B*A*B]
ans =
     0     0     0     0     0     0
     0     0     0     0     0     0
     1     0     0     0     0     0
>> [B*A B*A*B*A]
ans =
     0     1     0     0     0     1
     0     0     1     0     0     0
     0     0     0     0     0     0
```

9. It suffices to check that

$$f(X+E) = f(X) + \sum_{i=1}^{\infty} a_i \sum_{j=1}^{i} X^{j-1} E X^{i-j} + O(\|E\|^2),$$

since $L(X,E)$ is the linear term in this expansion. The matrix power series has the same radius of convergence as the given scalar series, so if $\|X\| < r$ we can scale $E \to \theta E$ so that $\|X + \theta E\| < r$ and the expansion is valid. But $L(X, \theta E) = \theta L(X,E)$, so the scaled expansion yields $L(X,E)$. $K(X)$ is obtained by using $\mathrm{vec}(AXB) = (B^T \otimes A)\mathrm{vec}(X)$.

10. We have $\mathrm{sign}(A) = A(A^2)^{-1/2} = A \cdot I^{-1/2} = A$.

11. $\operatorname{sign}(A) = \operatorname{sign}(A^{-1})$. The easiest way to see this is from the Newton iteration, because both $X_0 = A$ and $X_0 = A^{-1}$ lead to $X_1 = \frac{1}{2}(A + A^{-1})$ and hence the same sequence $\{X_k\}_{k\geq 1}$.

12. No: A^2 differs from I in the (1,3) and (2,3) entries. A quick way to arrive at the answer without computing A^2 is to note that if A is the sign of some matrix then since $a_{22} = a_{33} = 1$ we must have $A(2\colon 3, 2\colon 3) = I$, which is a contradiction.

16. $f(z) = 1/z$ for $z \neq 0$ and $f^{(j)}(0) = 0$ for all j. Hence f is discontinuous at zero. Nevertheless, $f(A)$ is defined.