Math 3607: Homework 7

Selected Solutions

2. (a) Let $X \in \mathbb{R}^{k\times k}$ be written as $X = VDV^{-1}$ where

$$V = \left[\begin{array}{c|c|c} \mathbf{v}_1 & \cdots & \mathbf{v}_k \end{array}\right] \in \mathbb{R}^{k\times k}, \qquad \text{(eigenvectors)}$$

and

$$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix} \in \mathbb{R}^{k\times k}. \qquad \text{(eigenvalues)}$$

Note that, for any $1 \le j \le n$,

$$X^j = \underbrace{\left(VDV^{-1}\right)\cdots\left(VDV^{-1}\right)}_{j \text{ copies}} = VD^jV^{-1},$$

and so

$$\begin{aligned} p(X) &= c_1 I + c_2 VDV^{-1} + c_3 VD^2V^{-1} + \cdots + c_n VD^{n-1}V^{-1} \\ &= V\underbrace{\left(c_1 I + c_2 D + c_3 D^2 + \cdots + c_n D^{n-1}\right)}_{=p(D)} V^{-1} = Vp(D)V^{-1}. \end{aligned}$$

Observe that $p(D)$ is a diagonal matrix whose (j, j)-entry is given by

$$c_1 + c_2\lambda_j + c_3\lambda_j^2 + \cdots + c_n\lambda_j^{n-1} = p(\lambda_j), \quad 1 \le j \le k.$$

That is,

$$p(D) = \begin{bmatrix} p(\lambda_1) & & \\ & \ddots & \\ & & p(\lambda_k) \end{bmatrix},$$

and so it only requires evaluations of p at the eigenvalues. Once it is constructed, one can form $p(A)$ by multiplying it by V and V^{-1}, two matrix multiplications.

(b) Since Horner's method is used in all scenarios, let's write it as a separate function.

```
function y = mypolyval(c, x)
%MYPOLYVAL evaluates a polynomial at points  x  given its coeffs.
% Input:
%   c   coefficient vector (c_1, c_2, ..., c_n)^T
%   x   points of evaluation
%       - if x is a scalar or a vector, use Horner's method
```

```
%         - if x is a square matrix, use the result from (a)
%         - otherwise, produce an error message.
    [k,m] = size(x);
    if k==1 || m==1
        y = horner(c, x);
    elseif k==m
        [V,D] = eig(x);
        y = V*diag(horner(c, diag(D)))/V;  % implementing part (a)
    else
        error('Input x must be a scalar, a vector, or a square
            matrix');
    end
end

function y = horner(c, x)
%HORNER Horner's method to evaluate polynomial
% Input:
%   c   coefficient vector (c_1, c_2, ..., c_n)^T
%   x   points of evaluation (either a scalar or a vector)
    n = length(c);
    y = c(n);
    for j = n-1:-1:1
        y = y.*x + c(j);
    end
end
```

Notes.

- Note that $p(D)$ is computed by `diag(horner(c, diag(D)))`, which entails three steps:
  i. Extraction of the eigenvalues of X into a column vector – the innermost expression, `diag(D)`.
  ii. Evaluation of p at the eigenvalues of X using Horner's method – the middle expression, `horner(....)`.
  iii. Construction of a diagonal matrix $p(D)$ – the outermost expression – `diag(....)`.

  This elegant construction relies on the dual functionality of the `diag` function, one for extraction of diagonal elements of a matrix and the other for construction of a diagonal matrix out of a vector.
- Pay attention to how $Vp(D)V^{-1}$ is implemented in MATLAB. In particular, as demonstrated in lecture, the right-multiplication by V^{-1} is done efficiently using the *forward slash* / rather than the *backward slash* \.

3. Recall that the nonzero singular values of A are the square roots of the nonzero eigenvalues of $A^{\mathrm{T}}A$[1]. So first compute $A^{\mathrm{T}}A$:

$$A^{\mathrm{T}}A = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

[1] The square roots of the nonzero eigenvalues of $AA^{\mathrm{T}} \in \mathbb{R}^{4\times 4}$ are also nonzero singular values of A. However, since the problem asks to solve a 2×2 problem, $A^{\mathrm{T}}A$ must be used as in this solution.

Then find its eigenvalues:

$$\left\{\begin{aligned} \det\left(\lambda I - A^{\mathrm{T}}A\right) &= \begin{vmatrix} \lambda-2 & -1 \\ -1 & \lambda-2 \end{vmatrix} \\ &= (\lambda-2)^2 - 1 \\ &= \lambda^2 - 4\lambda + 3. \end{aligned}\right\} \implies \lambda_1 = 3,\ \lambda_2 = 1.$$

Note that all eigenvalues are nonnegative; they are arranged in descending order so that the singular values are ordered properly. Hence, the two singular values of A are

$$\begin{aligned} \sigma_1 &= \sqrt{\lambda_1} = \sqrt{3}, \\ \sigma_2 &= \sqrt{\lambda_2} = 1. \end{aligned}$$

4. (a) Suppose that $A = U\Sigma V^{\mathrm{T}}$ is an SVD of A. Then

$$A^{\mathrm{T}} = \left(U\Sigma V^{\mathrm{T}}\right)^{\mathrm{T}} = V\Sigma^{\mathrm{T}}U^{\mathrm{T}} = V\Sigma U^{\mathrm{T}}.$$

Note that $\Sigma^{\mathrm{T}} = \Sigma$ since it is an $(n \times n)$ diagonal matrix. Since U and V are orthogonal matrices, the last factorization is an SVD of A^{T}. In particular, the singular values of A^{T} are the diagonal entries of Σ which are also the singular values of A.

(b) From the previous part, we know that both matrices share the same set of singular values. Since the 2-norm of a matrix is its largest singular value, it follows that $\|A\|_2 = \left\|A^{\mathrm{T}}\right\|_2$.