CSL361 Problem set 9: Eigenvalue and SVD Computation

April 17, 2017

1. (**Shift**) Show that if $Ax = \lambda x$ and σ is any scalar which is not an eigenvalue of A, then $(A - \sigma I)x = (\lambda - \sigma)x$. Thus the eigenvalues of $(A - \sigma I)$ are *shifted* from those of A by σ and the eigenvectors are unchanged.

2. (**Inverse**) Show that if A is nonsingular and $Ax = \lambda x$ with $x \neq 0$, then λ is necessarily nonzero, and $A^{-1}x = (1/\lambda)x$.

3. (**Power**) Show that if $Ax = \lambda x$ then $A^2 x = \lambda^2 x$. More generally, if k is any positive integer, then $A^k x = \lambda^k x$.

4. Given the *Rayleigh quotient* of a vector $x \in \mathbb{R}^m$:

$$r(x) = \frac{x^t AX}{x^t x}$$

show that the gradient of $r(x)$ (vector of partial derivatives with respect to coordinates x_j) is given as

$$\nabla r(x) = \frac{2}{x^t x}(Ax - r(x)x)$$

Conclude that at an eigenvector x of A, the gradient of $r(x)$ is the zero vector. Conversely, if $\nabla r(x) = 0$ with $x \neq 0$, then x is an eigenvector and $r(x)$ is the corresponding eigenvalue.

5. Let q_J be an eigenvector of A. From the fact that $\nabla r(q_J) = 0$ together with smoothness of the function $r(x)$ (everywhere except at the origin), conclude that

$$r(x) - r(q_J) = O(\|x - q_J\|^2) \text{ as } x \to q_J$$

6. Consider the *power* iteration:

   $v^{(0)}$ = some vector with $\|v^{(0)}\| = 1$
   for $k = 1, 2, \ldots$
   $\quad w = Av^{(k-1)}$
   $\quad v^{(k)} = w/\|w\|$
   $\quad \lambda^{(k)} = (v^{(k)})^t A v^{(k)}$

   Suppose that $|\lambda_1| > |\lambda_2| \geq \ldots |\lambda_m|$ and $q_1^t v^{(0)} \neq 0$. Show that the iterates satisfy

   $$\|v^{(k)} - (\pm q_1)\| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right) \text{ and } |\lambda^{(k)} - \lambda_1| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right)$$

   as $k \to \infty$.

7. Consider the *inverse* iteration:

   $v^{(0)}$ = some vector with $\|v^{(0)}\| = 1$
   for $k = 1, 2, \ldots$
   $\quad$ Solve $(A - \sigma I)w = v^{(k-1)}$ for w
   $\quad v^{(k)} = w/\|w\|$
   $\quad \lambda^{(k)} = (v^{(k)})^t A v^{(k)}$

   Suppose λ_J is the closest eigenvalue to σ and λ_K is the second closest, that is, $|\sigma - \lambda_J| < |\sigma - \lambda_K| \leq |\sigma - \lambda_j|$ for each $j \neq J$ and $q_J^t v^{(0)} \neq 0$. Then, show that the iterates of the inverse iteration satisfy

   $$\|v^{(k)} - (\pm q_J)\| = O\left(\left|\frac{\sigma - \lambda_J}{\sigma - \lambda_K}\right|^k\right) \text{ and } |\lambda^{(k)} - \lambda_J| = O\left(\left|\frac{\sigma - \lambda_J}{\sigma - \lambda_K}\right|^{2k}\right)$$

   as $k \to \infty$.

8. Consider the *Rayleigh quotient* iteration:

   $v^{(0)}$ = some vector with $\|v^{(0)}\| = 1$
   $\lambda^{(0)} = (v^{(0)})^t A v^{(0)}$
   for $k = 1, 2, \ldots$
   $\quad$ Solve $(A - \lambda^{(k-1)} I)w = v^{(k-1)}$ for w
   $\quad v^{(k)} = w/\|w\|$
   $\quad \lambda^{(k)} = (v^{(k)})^t A v^{(k)}$

Suppose λ_J is an eigenvalue A and $v^{(0)}$ is sufficiently close to q_J. Then, argue that for almost all starting vectors the iterates of the *Rayleigh quotient* iteration satisfy

$$\|v^{(k+1)} - (\pm q_J)\| = O(\|v^{(k)} - (\pm q_J)\|^3) \text{ and } |\lambda^{(k+1)} - \lambda_J| = O(|\lambda^{(k)} - \lambda_J|^3)$$

as $k \to \infty$.

9. (**Deflation**). Let H be any nonsingular matrix such that $Hx = \alpha e_1$, a scalar multiple of the first column of the identity matrix (*Householder* is a good choice for H). Show that the similarity transformation determined by H transforms A to a block triangular form

$$HAH^{-1} = \begin{bmatrix} \lambda_1 & b^t \\ 0 & B \end{bmatrix}$$

where B is a matrix of order $m-1$ having eigenvalues $\lambda_2, \ldots, \lambda_m$. Moreover if y_2 is an eigenvector of B corresponding to λ_2, then

$$x_2 = H^{-1} \begin{bmatrix} \alpha \\ y_2 \end{bmatrix} \text{ where } \alpha = \frac{b^t y_2}{\lambda_2 - \lambda_1}$$

is an eigenvector corresponding to λ_2 for the original matrix A, provided $\lambda_1 \neq \lambda_2$.

Conclude, that using deflation it is possible to determine all eigenvalues and eigenvectors of a matrix with any variation of the power iteration.

10. Consider the following algorithm known as *simultaneous iteration*:

$V^{(0)} =$ some arbitrary $m \times n$ matrix of rank n
for $k = 1, 2, \ldots$
    $V^{(k)} = AV^{(k-1)}$
$\hat{Q}^{(k)} \hat{R}^{(k)} = V^{(k)}$

Let $S_0 = \text{span}(V^{(0)})$ and let S be the invariant subspace spanned by the eigenvectors $x_1, x_2, \ldots, x_n$ of A corresponding to the n largest eigenvalues. Suppose that no non-zero vector in S is orthogonal to S_0. Show that for any $k > 0$, the columns of $V^{(k)}$ form a basis for $S_k = A^k S_0$, and, provided $\lambda_n > \lambda_{n+1}$, S_k converges to S (proof analogous to *power iteration*). Hence the final $\hat{Q}^{(k)}$ gives an orthogonal basis for the invariant subspace.

However, argue that the simultaneous iteration has the effect of carrying out power iteration of each column of $V^{(0)}$ and hence each column tends to converge to a multiple of the dominant eigenvector of A. Hence, the columns of $V^{(k)}$ form an increasingly ill-conditioned basis for S_k.

11. A remedy to the above is known as *orthogonal iteration*:

    $V^{(0)} =$ some arbitrary $m \times n$ matrix of rank n
    for $k = 1, 2, \ldots$
        $\hat{Q}^{(k)}\hat{R}^{(k)} = V^{(k-1)}$ (reduced QR factorization)
        $V^{(k)} = A\hat{Q}^{(k)}$

    where instead of orthogonalizing at the end, we orthogonalize at every iteration.

    Argue that that the matrices $V^{(k)}$ produced by the orthogonal version of simultaneous iteration converge to an $m \times n$ matrix V whose columns form a basis for same invariant subspace. Also, because $\text{span}(\hat{Q}^{(k)}) = \text{span}(V^{(k-1)})$, the matrices $\hat{Q}^{(k)}$ converge to an orthonormal basis for the same subspace.

    Also, we know that there exists an $n \times n$ matrix B such that $A\hat{Q} = \hat{Q}B$. Argue that for any j, $1 \leq j \leq n$, the first j columns of $\hat{Q}$ (or V) are the same as if the iteration has been carried out on the first j columns of A, and the remaining $n - j$ columns of $\hat{Q}$ can be expanded into a basis for the complementary subspace. Thus, if $\lambda_j > \lambda_{j+1}$ for $j = 1, \ldots, n$, then B must be triangular. Conclude that simultaneous orthogonal iterations lead to a *Schur decomposition* of A.

12. Consider the following iterations

    (a) Simultaneous orthogonal iteration

    $\underline{Q}^{(0)} = I$
    for $k = 1, 2, \ldots$
        $Z = A\underline{Q}^{(k-1)}$
        $\underline{Q}^{(k)}R^{(k)} = Z$ (QR factorization)
        $A^{(k)} = (\underline{Q}^{(k)})^t A \underline{Q}^{(k)}$

    (b) Unshifted QR iteration

$A^{(0)} = A$
for $k = 1, 2, \ldots$
  $Q^{(k)} R^{(k)} = A^{(k-1)}$ (QR factorization)
  $A^{(k)} = R^{(k)} Q^{(k)}$
  $\underline{Q}^{(k)} = Q^{(1)} Q^{(2)} \ldots Q^{(k)}$

Additionally, for both algorithms, let

$$\underline{R}^{(k)} = R^{(k)} R^{(k-1)} \ldots R^{(1)}$$

Show, by induction on k, that both generate identical sequences of matrices $\underline{R}^{(k)}$, $\underline{Q}^{(k)}$ and $A^{(k)}$, namely, those defined by the QR factorization of the k^{th} power of A,

$$A^k = \underline{Q}^{(k)} \underline{R}^{(k)}$$

together with the projection

$$A^{(k)} = (\underline{Q}^{(k)})^t A \underline{Q}^{(k)}$$

13.

14. Using all of the above convince yourself of the rationale behind the *practical QR algorithm*

$(Q^{(0)})^t A^{(0)} Q^{(0)} = A$ (Hessenberg reduction)
for $k = 1, 2, \ldots$
  Pick a shift $\mu^{(k)}$ (e.g., choose $\mu^{(k)} = A_{mm}^{(k-1)}$)
  $Q^{(k)} R^{(k)} = A^{(k-1)} - \mu^{(k)} I$ (QR factorization)
  $A^{(k)} = R^{(k)} Q^{(k)} + \mu^{(k)} I$ (re-combine factors in reverse order)
  If any sub-diagonal entry in $A^{(k)}$ is sufficiently close to zer, set it to zero to obtain
  $A^{(k)} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$ (deflation)
  and apply the QR algorithm to A_{11} and A_{22}