APPM 4/5720, 01-25-2016 Notes

1 Subspaces and Rank-Deficient Matrices

Let A be an $m \times n$ matrix of rank $k < \min(m, n)$.

- $A : X \to Y$, $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$ with decomposition $X = X_1 \oplus X_2$, $Y = Y_1 \oplus Y_2$.
- $X_1 = \text{row}(A)$, $X_2 = \text{null}(A) = X_1^\perp$ and $Y_1 = \text{col}(A)$, $Y_2 = Y_1^\perp$
- $\dim(X_1) = \dim(Y_1) = k$, $\dim(X_2) = n - k$, and $\dim(Y_2) = m - k$
- X_1 is the space spanned by the rows of A. Thus, $Ax = 0 \iff x \perp$ every row of A.
- Given $x \in X$, we can write $x = x_1 + x_2$ with $x_1 \in X_1$, $x_2 \in X_2$. Then, $Ax = Ax_1 + Ax_2 = Ax_1$ because $x_2 \in \text{null}(A)$. Therefore, A acts only on x_1 and A maps X_1 to Y_1.

Suppose $y \in Y$ is given where $y = y_1 + y_2$. Consider the equation

$$Ax = y \tag{1}$$

(1) has a solution $\iff y \in Y_1$ (so $y_2 = 0$). If z is a solution to $Az = y$ then $z + u$ is another solution for any $u \in X_2$.

1.1 Connection to SVD

Let $A = U_1 D_1 V_1^*$ where A is $m \times n$, U_1 is $m \times k$, D_1 is $k \times k$ and V_1^* is $k \times n$.

- The columns of U_1 form an orthonormal basis for Y_1.
- The columns of V_1 form an orthonormal basis for X_1.
- Given $x \in X$, set $x_1 = V_1 V_1^* x$ where $V_1 V_1^*$ is the projection operator onto the subspace X_1.
- Similarly, $x_2 = x - x_1 = (I - V_1 V_1^*)x$ where $(I - V_1 V_1^*)$ is the projection operator onto X_2.
- Given $y \in Y$, set $y_1 = U_1 U_1^* y$, then $y_2 = y - y_1 = (I - U_1 U_1^*)y$ where $U_1 U_1^*$ and $(I - U_1 U_1^*)$ are the projection operators onto Y_1 and Y_2 respectively.

Consider the full SVD

$$A = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} D_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix} \tag{2}$$

and equation (1). Then, $Ax = UDV^*x = y$. Since U is unitary (or $U^*U = I$) we have $DV^*x = U^*y$ so

$$\begin{bmatrix} D_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^*x \\ V_2^*x \end{bmatrix} = \begin{bmatrix} U_1^*y \\ U_2^*y \end{bmatrix}$$

This leads to the system of equations:

$$\begin{cases} U_1^*y = D_1V_1^*x \\ U_2^*y = 0 \end{cases} \tag{3}$$

From (3), we know there exists a solution $\iff U_2^*y = 0 \iff y \in Y_2$.

Since D is $k \times k$ (recall A is rank $k < \min(m, n)$), D is invertible and

$$V_1^*x = D_1^{-1}U_1^*y.$$

The Least Squares Solution is then $x = V_1D_1^{-1}U_1^*y$.

1.1.1 Least Squares Solution

Set $\hat{x} = V_1D_1^{-1}U_1^*y$. Then

1. $||A\hat{x} - y|| = ||U_2^*y|| = ||y_2|| = \inf_{x \in \mathbb{R}^n} ||Ax - y||$.
2. Among all $x \in \mathbb{R}^n$ such that $||Ax - y||$ is minimal, $\hat{x}$ is the shortest vector.

The matrix $A^\dagger = V_1D_1^{-1}U_1^*$ is the "Moore-Penrose" Inverse. If A is square and full rank, then $A^\dagger = A^{-1}$.

Consider the QR Factorization

$$AP = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix}. \tag{4}$$

Then

$$Ax = y \iff QRP^*x = y \iff RP^* = Q^*y \iff \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_p \\ x_n \end{bmatrix} = \begin{bmatrix} Q_1^*y \\ Q_2^*y \end{bmatrix},$$

where x_p are the k-number of pivoting coefficients and x_n are the $n - k$ "non-pivoting" coefficients corresponding to the product P^*x. From the above equivalence statements, we have $\text{col}(A) = \text{col}(Q_1)$ and

$$\inf_{x \in \mathbb{R}^n} ||Ax - y|| = ||Q_2^*y|| = ||y_2|| = ||(I - Q_1Q_1^*)y||$$

which minimizes the residual. Furthermore, $R_{11}x_p + R_{12}x_n = Q_1^* y$. Setting $x_p = R_{11}^{-1} Q_1^* y$ and $x_n = 0$, we have

$$x = P \begin{bmatrix} x_p \\ 0 \end{bmatrix}$$

which is the most sparse solution, NOT necessarily the shortest.

2 The "Two-Stage" Approach to Low Rank Approximation

Let A be a $m \times n$ matrix of rank k. Suppose Q is an $m \times k$ matrix whose columns form an orthonormal basis for $\text{col}(A)$. We can then easily compute an SVD of A:

1. $A = QQ^*A$ since $\text{col}(A) = \text{col}(Q)$.
2. Set $B = Q^*A$, then $A = QB$ where B is $k \times n$ and Q is $m \times k$ to get a low rank factor of A.
3. Compute the full SVD of B, so $A = QB = Q\hat{U}DV^*$.
4. Let $Q\hat{U} = U$ (because the product of two unitary matrices is a unitary matrix). Then A $= UDV^*$ and we have the SVD of A.

We needed only $\approx k^2(m+n)$ FLOPS, and a matrix-matrix product $B = Q^*A$.

Question: How do we find Q?