Lec 8: Matrix Decomposition

Ailin Zhang

Agenda

- Matrix Decomposition (QR)
- Gram-Schmidt Process
- Householder Reflection

Matrix Decomposition

QR decomposition decomposes a matrix $\mathbf{X}$ into a product $\mathbf{X} = QR$ where Q is an orthogonal matrix and R is an upper triangular matrix. It does not require the computation of the cross-product matrix such as $\mathbf{X}^\top \mathbf{X}$.

Let $Q = (q_1, q_2, \ldots, q_n)$ be an orthogonal matrix, $Q^\top Q = QQ^\top = I$, then Q forms an orthogonal basis:

(1) For each vector q_i, $\|q_i\| = 1$.

(2) For any two different vectors q_i and q_j, $\langle q_i, q_j \rangle = 0$, i.e., $q_i \perp q_j$.

For any vector v, we have

(1) Analysis: $u_i = \langle v, q_i \rangle = q_i^\top v$ is the coordinate of v on the axis q_i, for $i = 1, ..., n$, i.e., $u = Q^\top v$.

(2) Synthesis: $v = \sum_{i=1}^n q_i u_i = Qu$.

From (1) and (2), Q and $Q^\top$ are inverse of each other.

Gram-Schmidt Process

Let $W = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & -3 \\ 1 & 4 & -4 \\ 1 & 5 & -2 \end{bmatrix}$

Find the orthogonal basis for W

Gram-Schmidt Process

$$u_1 = a_1, \qquad q_1 = \frac{u_1}{\|u_1\|_2}, \quad \tilde{r}_{11} = \frac{1}{\|u_1\|_2}$$

$$u_2 = a_2 - (q_1^T a_2)q_1, \qquad q_2 = \frac{u_2}{\|u_2\|_2}, \quad \tilde{r}_{12} = -\frac{q_1^T a_2}{\|u_2\|_2}, \quad \tilde{r}_{22} = \frac{1}{\|u_2\|_2}$$

$$\vdots$$

$$u_n = a_n - \sum_{j=1}^{n-1}(q_j^T a_n)q_j, \qquad q_n = \frac{u_n}{\|u_n\|_2}, \quad \tilde{r}_{jn} = -\frac{q_j^T a_n}{\|u_n\|_2}, \quad \tilde{r}_{nn} = \frac{1}{\|u_n\|_2}$$

Gram-Schmidt Process

$$A \begin{bmatrix} \tilde{r}_{11} & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & \tilde{r}_{12} & & \\ & \tilde{r}_{22} & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & & & \tilde{r}_{1n} \\ & 1 & & \tilde{r}_{2n} \\ & & \ddots & \vdots \\ & & & \tilde{r}_{n1} \end{bmatrix} = Q$$

Projection Matrix

In Gram-Schmidt, we can treat each vector q_j as a product of projection of A_j,where the projection matrix is $P_j = I - Q_{j-1}Q_{j-1}^T$ and it satisfies the following:

1. $P^2 = P$
2. $I - P$ is also a projection matrix

Gram-Schmidt Process with Projection Matrix

$$Q_j = \begin{bmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_j \\ | & | & & | \end{bmatrix}$$

$$\begin{aligned} q_1 &= \frac{P_1 a_1}{||P_1 a_1||} \\ q_2 &= \frac{P_2 a_2}{||P_2 a_2||} \\ &\vdots \\ q_n &= \frac{P_n a_n}{||P_n a_n||} \end{aligned}$$

$$P_1 = I$$

Exercise

For QR decomposition, $R = Q^T A$

Find Q and R for $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$

Householder reflection

To obtain a QR decomposition, we can apply the Householder reflections repeatedly. Given an $n \times p$ matrix $\mathbf{X}$, as the first step, we want to find an orthogonal transformation H_1 such that only the first element in the first column is non-zero after the transformation:

$$\begin{bmatrix} x_{11} & x_{12} & \dots & y_1 \\ x_{21} & x_{22} & \dots & y_2 \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & y_n \end{bmatrix} \xrightarrow{H_1} \begin{bmatrix} x_{11}^* & x_{12}^* & \dots & y_1^* \\ 0 & x_{22}^* & \dots & y_2^* \\ \dots & \dots & \dots & \dots \\ 0 & x_{n2}^* & \dots & y_n^* \end{bmatrix}$$

Graphical Representation

Since the orthogonal transformation preserves the length of vectors, we know

$$|\mathbf{X}_1^*| = |\mathbf{X}_1| = \sqrt{x_{11}^2 + x_{12}^2 + \ldots + x_{1n}^2},$$

which means the value of x_{11^*} is determined by

$$x_{11}^* = \pm |\mathbf{X}_1| .$$

The sign of x_{11}^* is chosen as the opposite of x_{11} for numerical stability.

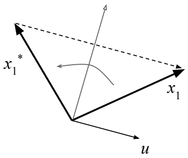

Figure 1: Household Reflection

Discussion: Choice of X_{11}^*

The sign of x_{11}^* is chosen as the opposite of x_{11} for numerical stability.

Discussion: Recursion

$$\begin{bmatrix} x_{11}^* & x_{12}^* & \dots & y_1^* \\ 0 & x_{22}^* & \dots & y_2^* \\ \dots & \dots & \dots & \dots \\ 0 & x_{n2}^* & \dots & y_n^* \end{bmatrix} \xrightarrow{H_2} \begin{bmatrix} x_{11}^* & x_{12}^* & \dots & y_1^* \\ 0 & x_{22}^* & \dots & y_2^* \\ \dots & 0 & \dots & \dots \\ 0 & 0 & \dots & y_n^* \end{bmatrix}$$

Derivation

To find a transformation H which can rotate the vector $\mathbf{X}_1$ to $\mathbf{X}_1^*$, one simple way is to construct an isosceles triangle where $\mathbf{X}_1^*$ is a reflection of $\mathbf{X}_1$:

$$\mathbf{X}_1^* = \mathbf{X}_1 - 2\langle \mathbf{X}_1, u\rangle u = \mathbf{X}_1 - 2uu^\top \mathbf{X}_1 = H_1 \mathbf{X}_1,$$

where

$$u = \frac{\mathbf{X}_1 - \mathbf{X}_1^*}{|\mathbf{X}_1 - \mathbf{X}_1^*|}, \quad H_1 = I - 2uu^\top.$$

Recursion:

Let $\mathbf{X}^{(1)}$ be the sub-matrix of the resulting $\mathbf{X}$ with the first row and the first column removed from $\mathbf{X}$. Apply the Householder reflection on the sub-matrix $\mathbf{X}^{(1)}$, while maintaining the first row and the first column of $\mathbf{X}$. This amounts to left multiplying $\mathbf{X}$ by an orthogonal matrix H_2.

Let $H = H_p \dots H_2 H_1$, we have $H\mathbf{X} = R$. Let $Q = H^\top$, we obtain the QR decomposition $\mathbf{X} = QR$

R Code

```
myQR <- function(A) {
    n <- nrow(A)
    m <- ncol(A)
    R <- A
    Q <- diag(n)
    # Perform Householder reflection one column at a time (for all but last column)
for (k in 1:(m - 1)) {
#x,g,v,s just break up the calculation of u
x <- matrix(rep(0, n), nrow = n)
x[k:n, 1] <- R[k:n, k]
g <- sqrt(sum(x^2))
v <- x
v[k] <- x[k] + sign(x[k,1]) * g
s = sqrt(sum(v^2))
if (s != 0) {
            u <- v / s
            R <- R - 2 * u %*% t(u) %*% R
            Q <- Q - 2 * u %*% t(u) %*% Q
} }
    result <- list(Q=t(Q), R=R)
return(result) }
```

Python Code

```
import numpy as np
from scipy import linalg

def qr(A):
    n, m = A.shape
    R = A.copy()
    Q = np.eye(n)

    for k in range(m-1):
      x = np.zeros((n, 1))
      x[k:, 0] = R[k:, k]
      v=x
      v[k] = x[k] + np.sign(x[k,0]) * np.linalg.norm(x)
      s = np.linalg.norm(v)
      if s != 0:
        u=v/s
        R -= 2 * np.dot(u, np.dot(u.T, R))
        Q -= 2 * np.dot(u, np.dot(u.T, Q))
    Q = Q.T
    return Q, R

A = rand(100,100)
B = copy(A)
(Q,R) = householder(A)
B = Q*R
```

Gram-Schmidt Vs Householder

1. Column vs row transformation
2. Gram-Schmidt suffers more from numerical instability

Application of QR: Linear regression

We rotate the matrix $(\mathbf{XY})$ by QR decomposition, by applying the Householder reflections for $j = 1, \ldots, p$,

$$\begin{bmatrix} \mathbf{X} & \mathbf{Y} \end{bmatrix} \xrightarrow{Q^\top} \begin{bmatrix} R & \mathbf{Y}^* \end{bmatrix} = \begin{bmatrix} R_1 & \mathbf{Y}_1^* \\ 0 & \mathbf{Y}_2^* \end{bmatrix}$$

where R_1 is a upper triangular squared matrix. To solve the least squares problem,

$$\min_\beta \|\mathbf{Y}^* - R\beta\|^2 = \min_\beta \left(\|Y_1^* - R_1\beta\|^2 + \|\mathbf{Y}_2^*\|^2 \right)$$

the solution $\hat{\beta} = R_1^{-1}\mathbf{Y}_1^*$ and $\mathrm{RSS} = \|\mathbf{Y}_2^*\|^2$. Since R_1 is an upper triangular matrix, we can solve the elements of $\hat{\beta}$ in reverse order $\hat{\beta}_p, \hat{\beta}_{p-1}, \ldots, \hat{\beta}_1$. It is numerically stable and efficient.

Application of QR: Approximating Eigenvalues of A

- Find $A = Q_0 R_0$
- Let $A_1 = R_0 Q_0$ Find $A_1 = Q_1 R_1$
- Let $A_2 = R_1 Q_1$ Find $A_2 = Q_2 R_2$
- etc...

1. Fact 1: A_k has the same eigenvalues as A for $k = 1, 2, 3, \ldots$
2. Fact 2: If the eigenvalues of A are real, with distinct absolute values then Ak becomes upper triangular as $k \to \infty$

Conclusion: The diagonal entries of A_k approximate the eigenvalues of A