

MAS VIMU Matematičke osnove mašinskog učenja

Zoran Ovcin

2026-01-06

3.8.3 Gram-Schmidt Orthogonalization MML-Book

Neka je V vektorski prostor sa unutrašnjim proizvodom $\langle \cdot, \cdot \rangle$.

Vektori $\mathbf{u}, \mathbf{v} \in V$ su ortogonalni ako je $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Neka je $(\mathbf{u}_1, \dots, \mathbf{u}_k)$ uređena k-torka ortogonalnih vektora iz V . Onda:

- Vektori $\mathbf{u}_1, \dots, \mathbf{u}_k$ su linearno nezavisni.
- Neka je $U = \text{span}[\mathbf{u}_1, \dots, \mathbf{u}_k]$ i neka je $\mathbf{w} := \sum_{j=1}^k \frac{\langle \mathbf{v}, \mathbf{u}_j \rangle}{\langle \mathbf{u}_j, \mathbf{u}_j \rangle} \mathbf{u}_j$. Onda je $\mathbf{v} - \mathbf{w}$ ortogonalan na sve vektore prostora U ($\mathbf{u} \in U \Rightarrow \langle \mathbf{w}, \mathbf{u} \rangle = 0$).
- Kažemo da je $\mathbf{w} = \pi_U(\mathbf{v}) \in U$ ortogonalna projekcija vektora $\mathbf{v} \in V$ na potprostor U .
- Skalare $\frac{\langle \mathbf{v}, \mathbf{u}_j \rangle}{\langle \mathbf{u}_j, \mathbf{u}_j \rangle}$ nazivamo Furijeovi koeficijenti.
- Furijeovi koeficijenti daju jedinstvenu reprezentaciju ortogonalne projekcije $\mathbf{w}$ vektora $\mathbf{v} \in V$ u bazi $(\mathbf{u}_1, \dots, \mathbf{u}_k)$.
- Od svih vektora $\mathbf{u} \in U$ udaljenost $\|\mathbf{w} - \mathbf{v}\| = \sqrt{\langle \mathbf{w} - \mathbf{v}, \mathbf{w} - \mathbf{v} \rangle}$ vektora $\mathbf{v}$ i $\mathbf{w}$ je najmanja.
- Ako $\mathbf{v} \in U$, onda je $\pi_U(\mathbf{v}) = \mathbf{v}$.

Neka je $(\mathbf{b}_1, \dots, \mathbf{b}_n)$ uređena baza prostora V . Neka je

$$\mathbf{u}_1 := \mathbf{b}_1$$

$$\mathbf{u}_k := \mathbf{b}_k - \pi_{\text{span}[\mathbf{u}_1, \dots, \mathbf{u}_{k-1}]}(\mathbf{b}_k), k = 2, \dots, n$$

Onda je $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ ortogonalna baza prostora V .

Primenjeno iterativno za $k = 1, \dots, n$, ovo je Gram-Schmidt potupak ortogonalizacije:

$$\mathbf{u}_1 := \mathbf{b}_1$$

$$\mathbf{u}_k := \mathbf{b}_k - \sum_{j=1}^{k-1} \frac{\langle \mathbf{b}_k, \mathbf{u}_j \rangle}{\langle \mathbf{u}_j, \mathbf{u}_j \rangle} \mathbf{u}_j, k = 2, \dots, n$$

Dobijene vektore možemo normirati $\mathbf{e}_k := \frac{1}{\sqrt{\langle \mathbf{u}_k, \mathbf{u}_k \rangle}} \mathbf{u}_k, k = 1, \dots, n$ i dobiti ortonormiranu bazu $(\mathbf{e}_1, \dots, \mathbf{e}_n)$.

Originalni vektori izraženi u dobijenoj ortonormiranoj bazi su: $\mathbf{b}_1 = \langle \mathbf{e}_1, \mathbf{b}_1 \rangle \mathbf{e}_1$

$$\mathbf{b}_2 = \langle \mathbf{e}_1, \mathbf{b}_2 \rangle \mathbf{e}_1 + \langle \mathbf{e}_2, \mathbf{b}_2 \rangle \mathbf{e}_2$$

$\vdots$

$$\mathbf{b}_n = \sum_{j=1}^n \langle \mathbf{e}_j, \mathbf{b}_n \rangle \mathbf{e}_j$$

Ako vektore baze $\mathbf{b}_k \in \mathbb{R}^m, k = 1, \dots, n$, smestimo u matricu $B = [\mathbf{b}_1, \dots, \mathbf{b}_n]$, onda $B = Q R$, gde je

$$Q = [\mathbf{e}_1, \dots, \mathbf{e}_n] \text{ i } R = \begin{bmatrix} \langle \mathbf{e}_1, \mathbf{b}_1 \rangle & \langle \mathbf{e}_1, \mathbf{b}_2 \rangle & \cdots & \langle \mathbf{e}_1, \mathbf{b}_n \rangle \\ 0 & \langle \mathbf{e}_2, \mathbf{b}_2 \rangle & \cdots & \langle \mathbf{e}_2, \mathbf{b}_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \langle \mathbf{e}_n, \mathbf{b}_n \rangle \end{bmatrix}$$

Pri tome je matrica Q ortogonalna: $Q^T Q = \mathbf{I}_n$. Ako je matrica Q kvadratna, onda je $Q^{-1} = Q^T$.

Ako je unutrašnji proizvod uobičajeni dot product, onda

$$\mathbf{u}_1 := \mathbf{b}_1$$

$$\mathbf{u}_k := \mathbf{b}_k - \sum_{j=1}^{k-1} \frac{\mathbf{b}_k^T \mathbf{u}_j}{\mathbf{u}_j^T \mathbf{u}_j} \mathbf{u}_j, k = 2, \dots, n.$$

Normiranje ortogonalne baze: $\mathbf{e}_k = \mathbf{u}_k / \|\mathbf{u}_k\| = \mathbf{u}_k / \sqrt{\mathbf{u}_k^T \mathbf{u}_k}, k = 1, \dots, n$.

Neka su kolone matrice $A = [\mathbf{a}_1, \dots, \mathbf{a}_n] \in \mathbb{R}^{m \times n}$ linearno nezavisne, odnosno, neka čine bazu potprostora $U = \text{span}[\mathbf{a}_1, \dots, \mathbf{a}_n] \subseteq \mathbb{R}^m$.

Onda sledeći algoritam kao izlaz daje $Q \in \mathbb{R}^{m \times n}$, matricu koja kao kolone sadrži vektore ortonormirane baze prostora U i gornje trougaonu matricu $R \in \mathbb{R}^{n \times n}$ tako da je $A = Q R$.

Ako se (u zadatoj toleranciji) neki od vektora matrice A pojavi kao linearna kombinacija prethodnih, onda algoritam staje, zadati vektori su linearno zavisni.

```
gramSchmidt = function(A, tol = 1e-12) {
  m = nrow(A)
  n = ncol(A)
  Q = matrix(0, m, n)
  R = matrix(0, n, n)
  for (j in 1:n) {
    v = A[,j]
    if (j>1){
      for (i in 1:(j - 1)) {
        R[i,j] = sum(Q[,i] * A[,j])
        # R[i,j] = sum(Conj(Q[,i]) * A[,j]) # Za kompleksne matrice
        v = v - R[i,j] * Q[,i]
      }
    }
    R[j,j] = sqrt(sum(v^2))
    # R[j, j] = sqrt(sum(Mod(v)^2)) # Za kompleksne matrice
    if (R[j,j] < tol) {
      stop("Kolone su linearno zavisne (sa datom tolerancijom).")
    }
    Q[,j] = v / R[j,j]
  }
  list(Q=Q,R=R)
}

A = matrix(c(1, 2, 4,
             2, 4, 5,
             4, 7, 5,
             1, 1, 6), ncol = 3, byrow = T)
```

```

rez = gramSchmidt(A)
rez$Q

##           [,1]      [,2]      [,3]
## [1,] 0.2132007 0.2445580 0.5362779
## [2,] 0.4264014 0.4891160 0.4731864
## [3,] 0.8528029 -0.0978232 -0.4942169
## [4,] 0.2132007 -0.8314972 0.4942169

rez$R

##           [,1]      [,2]      [,3]
## [1,] 4.690416 8.3148279 8.528029
## [2,] 0.000000 0.9293204 -2.054287
## [3,] 0.000000 0.0000000 5.005260

rez$Q %*% rez$R

##           [,1] [,2] [,3]
## [1,]      1      2      4
## [2,]      2      4      5
## [3,]      4      7      5
## [4,]      1      1      6

zapsmall(t(rez$Q) %*% rez$Q)

##           [,1] [,2] [,3]
## [1,]      1      0      0
## [2,]      0      1      0
## [3,]      0      0      1

A = matrix(c(1, 1, 1,
             1, 1+1e-11, 1,
             1, 1, 1+1e-11), ncol = 3, byrow = T)
rez = gramSchmidt(A)
t(rez$Q) %*% rez$Q

##           [,1]      [,2]      [,3]
## [1,] 1.000000e+00 -4.710277e-05 -4.710278e-05
## [2,] -4.710277e-05 1.000000e+00 1.000000e+00
## [3,] -4.710278e-05 1.000000e+00 1.000000e+00

zapsmall(t(rez$Q) %*% rez$Q)

##           [,1]      [,2]      [,3]
## [1,] 1.00e+00 -4.71e-05 -4.71e-05
## [2,] -4.71e-05 1.00e+00 1.00e+00
## [3,] -4.71e-05 1.00e+00 1.00e+00

```

QR dekompozicija matrice se koristi za kreiranje niza matrica koje konvergiraju ka Schurovoj formi (Schur form) matrice u kojoj se na dijagonali nalaze karakteristične vrednosti matrice sa sve višestrukostima.