7.9. Orthonormal basis and the Gram-Schmidt Process

We can find an orthonormal basis for any vector space using Gram-Schmidt process. Such bases are very useful.

Orthogonal projections can be computed using dot products

Fourier series, wavelets, and so on from these.

Orthogonal basis. Orthonormal basis

- Orthogonal basis: A basis that is an orthogonal set.
- Orthonormal basis: A basis that is an orthonrmal set.
- Example 1: {(0,1,0), (1,0,1), (-1,0,1)}
- Example 2: {(3/7,-6/7,2/7),(2/7,3/7,6/7), (6/7,2/7,-3/7)}
- Example 3: The standard basis of R^n.

Theorem 7.9.1 *An orthogonal set of nonzero vectors in R^n is linearly independent.*

- Proof: v_1,v_2,..,v_k Orthogonal set.
  - Suppose c_1v_1+c_2v_2+…+c_kv_k=0.
  - Dot with v_1. c_1v_1.v_1=0. Since v_1 has nonzero length, c_1=0.
  - Do for each v_js. Thus all c_j=0.
- Thus an orthogonal (orthonormal) set of n nonzero vectors is a basis always.

How to find these?

Orthogonal projections using orthonormal projections

- Proj_W x = $M(M^TM)^{-1}M^T(x)$.
- Recall M has columns that form a basis of W.
- Suppose we chose the orthonormal basis of W.
- $M^TM=I$ by orthonormality.
- Thus Proj_w(x)=MM^Tx.
- $P=MM^T$.
- Example 4.

Theorem 7.9.2

(*a*) *If* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *is an orthonormal basis for a subspace* W *of* R^n, *then the orthogonal projection of a vector* $\mathbf{x}$ *in* R^n *onto* W *can be expressed as*

$$\text{proj}_W \mathbf{x} = (\mathbf{x} \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{x} \cdot \mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{x} \cdot \mathbf{v}_k)\mathbf{v}_k \tag{7}$$

(*b*) *If* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *is an orthogonal basis for a subspace* W *of* R^n, *then the orthogonal projection of a vector* $\mathbf{x}$ *in* R^n *onto* W *can be expressed as*

$$\text{proj}_W \mathbf{x} = \frac{\mathbf{x} \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\mathbf{x} \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\mathbf{x} \cdot \mathbf{v}_k}{\|\mathbf{v}_k\|^2}\mathbf{v}_k \tag{8}$$

- Proof: (a) M=[v_1,v_2,..,v_k].

$$proj_W x = M(M^T x) = [v_1, v_2, \ldots, v_k]\begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_k^T \end{bmatrix} x$$

$$= [v_1, v_2, \ldots, v_k]\begin{bmatrix} v_1^T x \\ v_2^T x \\ \vdots \\ v_k^T x \end{bmatrix} = (x.v_1)v_1 + (x.v_2)v_2 + \ldots + (x.v_k)v_k$$

- Proof(b): Divide by the lengths to obtain an orthonormal basis of W. Apply (a).
- Note: Even if W=R^n, one can use the same formula.

Theorem 7.9.3 *If P is the standard matrix for an orthogonal projection of R^n onto a subspace of R^n, then* $\operatorname{tr}(P) = \operatorname{rank}(P)$.

- Proof: P=MM^T=v_1v_1^T+…+v_kv_k^T.
  - trP=tr(v_1v_1^T)+…+tr(v_kv_k^T)=v_1.v_1+…+v_k.v_k=k
  - This by Formula 27 in Sec 3.1.
- Example 7: 13/49+45/49+40/49=2 (Example 4)

Linear combinations of orthonormal basis vectors.

- If w is in W, then proj_W(w)=w. In particular, if W=R^n, and w any vector, we have

Theorem 7.9.4

(*a*) *If* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *is an orthonormal basis for a subspace* W *of* R^n, *and if* $\mathbf{w}$ *is a vector in* W, *then*

$$\mathbf{w} = (\mathbf{w} \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{w} \cdot \mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{w} \cdot \mathbf{v}_k)\mathbf{v}_k \tag{11}$$

(*b*) *If* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *is an orthogonal basis for a subspace* W *of* R^n, *and if* $\mathbf{w}$ *is a vector in* W, *then*

$$\mathbf{w} = \frac{\mathbf{w} \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\mathbf{w} \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\mathbf{w} \cdot \mathbf{v}_k}{\|\mathbf{v}_k\|^2}\mathbf{v}_k \tag{12}$$

- The above formula is very useful to find “coordinates” given an orthonormal basis.
- Example 8:

Gram-Schmidt orthogonalization process

- W a nonzero subspace $\{w_1, w_2, .., w_k\}$ Any basis
- We will produce orthogonal basis $\{v_1, v_2, .., v_k\}$
- Let $v_1 = w_1$.
- $v_2 = w_2 - proj_{w_1}(w_2) = w_1 - v_1 (w_2 . v_1)/||v_1||^2$.
  - v_2 is not zero. (Otherwise, $w_2 = proj_{w_1}(w_2)$. $w_1 // w_2$).
  - $\{v_1, v_2\}$ orthogonal set. Let $W_2 = Span\{v_1, v_2\}$
- $v_3 = w_3 - proj_{W_2}(w_3) = w_3 - v_1 (w_3 . v_1)/||v_1||^2 - v_2 (w_3 . v_2)/||v_2||^2$.
- v_3 is nonzero since w_3 is not in W_2 by independence of $\{w_1, w_2, w_3\}$. v_3 is orthogonal to v_1 and v_2.

- We obtained orthogonal set of $v_1, v_2, \ldots, v_l$. Let $W_l = \mathrm{Span}\{v_1, \ldots, v_l\}$.
- $v_{l+1} = w_{l+1} - \mathrm{proj}_{W_l}(w_{l+1}) = w_{l+1} - v_1 (w_{l+1} \cdot v_1)/||v_1||^2 - \ldots - v_l (w_{l+1} \cdot v_l)/||v_l||^2$
- Then v_{l+1} is not 0 since w_{l+1} is not in W_l.
- v_{l+1} is orthogonal to $v_1, .., v_l$.
    - $v_i \cdot (w_{l+1} - v_1 (w_{l+1} \cdot v_1)/||v_1||^2 - \ldots - v_l (w_{l+1} \cdot v_l)/||v_l||^2$
      $= v_i \cdot w_{l+1} - v_i \cdot v_i \, (w_{l+1} \cdot v_i)/||v_i||^2 = 0$ for $i = 1, .., l$.
- Finally, we achieve $v_1, v_2, .., v_k$.
- We can normalize to obtain an orthonormal basis.

- Example 9: (0,0,0,1),(0,0,1,1),(0,1,1,1),(1,1,1,1).
- Example 10: x+y+z+2t = 0, 2x+y+z+t=0.
- Properties:

Theorem 7.9.6 *If* $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ *is a basis for a nonzero subspace of* R^n, *and if* $S' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *is the corresponding orthogonal basis produced by the Gram–Schmidt process, then:*

(*a*) $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_j\}$ *is an orthogonal basis for* $\text{span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_j\}$ *at the* jth *step.*

(*b*) $\mathbf{v}_j$ *is orthogonal to* $\text{span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{j-1}\}$ *at the* jth *step* $(j \geq 2)$.

Extending the orthonormal set to orthonormal basis.

Theorem 7.9.7 *If W is a nonzero subspace of R^n, then:*

(*a*) *Every orthogonal set of nonzero vectors in W can be enlarged to an orthogonal basis for W.*

(*b*) *Every orthonormal set in W can be enlarged to an orthonormal basis for W.*

- Proof (a): Given v_1,..,v_k. Add v_k+1 orthogonal to Span{v_1,..,v_k}. Add v_k+2 orthogonal to Span{v_1,v_2,..,v_k,v_k+1}. By induction….
- Proof (b): see book