7.7. The projection theorem and its implications

Orthogonal projections, formula

Orthogonal projections to a line in R^2

- Let us obtain a formular for projection to a line containing a nonzero vector a.
- x=x_1+x_1, x_1=ka. x_2 is orthogonal to a.
- (x-ka).a=0. x.a-k(a.a)=0. $k=x.a/||a||^2$.
- $x\_1=(x.a/||a||^2)a$.
- $Proj\_a (x)=(x.a/||a||^2)a$.
- Example 1: The matrix expression.

Orthogonal projections onto a line through O in R^n.

Theorem 7.7.1 *If* $\mathbf{a}$ *is a nonzero vector in* R^n, *then every vector* $\mathbf{x}$ *in* R^n *can be expressed in exactly one way as*

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 \tag{6}$$

where $\mathbf{x}_1$ *is a scalar multiple of* $\mathbf{a}$ *and* $\mathbf{x}_2$ *is orthogonal to* $\mathbf{a}$ (*and hence to* $\mathbf{x}_1$). *The vectors* $\mathbf{x}_1$ *and* $\mathbf{x}_2$ *are given by the formulas*

$$\mathbf{x}_1 = \frac{\mathbf{x}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \quad \textit{and} \quad \mathbf{x}_2 = \mathbf{x} - \mathbf{x}_1 = \mathbf{x} - \frac{\mathbf{x}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \tag{7}$$

- Proof: Omit

Definition 7.7.2 If $\mathbf{a}$ is a nonzero vector in R^n, and if $\mathbf{x}$ is any vector in R^n, then the ***orthogonal projection of*** $\mathbf{x}$ ***onto*** $\text{span}\{\mathbf{a}\}$ is denoted by $\text{proj}_{\mathbf{a}}\mathbf{x}$ and is defined as

$$\text{proj}_{\mathbf{a}}\mathbf{x} = \frac{\mathbf{x}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \tag{11}$$

The vector $\text{proj}_{\mathbf{a}}\mathbf{x}$ is also called the ***vector component of*** $\mathbf{x}$ ***along*** $\mathbf{a}$, and $\mathbf{x} - \text{proj}_{\mathbf{a}}\mathbf{x}$ is called the ***vector component of*** $\mathbf{x}$ ***orthogonal to*** $\mathbf{a}$.

Projection operator in R^n

- $T(x)= proj\_a(x) = (x.a/||a||^2)a$
- Orthogonal projection of R^n onto span{a}.

Theorem 7.7.3 *If* $\mathbf{a}$ *is a nonzero vector in* R^n*, and if* $\mathbf{a}$ *is expressed in column form, then the standard matrix for the linear operator* $T(\mathbf{x}) = proj_{\mathbf{a}}\mathbf{x}$ *is*

$$P = \frac{1}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T \qquad (16)$$

This matrix is symmetric and has rank 1.

- Proof: $T(e\_j)=((e\_j.a)/||a||^2)a = (a\_j/||a||^2)a$.
  - $P=[a\_1a|a\_2a|\ldots|a\_na]/||a||^2=a[a\_1,a\_2,..,a\_n]/||a||^2=$
  - aa^T/a^Ta.

- If a is a unit vector u. Then $P=uu^T$.
- Example 4. P_θ again
- Example 5.

Projection theorem

Theorem 7.7.4 (***Projection Theorem for Subspaces***) *If W is a subspace of R^n, then every vector $\mathbf{x}$ in R^n can be expressed in exactly one way as*

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 \tag{20}$$

where $\mathbf{x}_1$ is in W and $\mathbf{x}_2$ is in $W^\perp$.

- Proof: Let {w_1,…,w_k} be the basis of W. Let M be the nxk matrix with columns w_1,w_2,..,w_k. k ≤n.
  - W column space of M. W^c null space of M^T.
  - Write x=x_1+x_2, x_1 in W and x_2 in W^c.
  - x_1=Mv and M^T(x_2)=0 or M^T(x-x_2)=0.
  - M^T(x-Mv)=0.
  - This has a unique solution <-> x_1,x_2 exist and are unique.

- Rewrite $M^TMv=M^Tx$.
  - M^TM is kxk-matrix
  - M has a full column rank as w_1,..,w_k are independent.
  - M^TM is invertible by Theorem 7.5.10.
  - $v=(M^TM)^{-1}M^Tx$.
- x = proj_W(x) + proj_W^c(x).
- Since proj_W(x)=x_1=Mv, we have

Theorem 7.7.5 *If W is a nonzero subspace of R^n, and if M is any matrix whose column vectors form a basis for W, then*

$$\text{proj}_W \mathbf{x} = M(M^TM)^{-1}M^T\mathbf{x} \tag{25}$$

for every column vector $\mathbf{x}$ in R^n.

- $T(x) = Proj\_W(x) = M(M^TM)^{-1}M^T(x)$.
- Matrix is $P = M(M^TM)^{-1}M^T$
- Orthogonal projection of R^n to W.
- This extends the previous formula.
- Example 6.

Condition for orthogonal projection

- $P^T=(M(M^TM)^{-1}M^T\ )^T=M((M^TM)^{-1})^TM^T = M(M^TM)^{-1}M^T=P.$
- $P^2=M(M^TM)^{-1}M^T\ (M(M^TM)^{-1}M^T = M(M^TM)^{-1}(M^TM)(M^TM)^{-1}M^T$ $=M(M^TM)^{-1}M^T=P.$
- $P^2=P.$

Theorem 7.7.6 *An $n \times n$ matrix P is the standard matrix for an orthogonal projection of R^n onto a k-dimensional subspace of R^n if and only if P is symmetric, idempotent, and has rank k. The subspace W is the column space of P.*

Strang diagrams

- Ax=b. A mxn matrix
- row(A), null(A) are orthogonal complements
- col(A) and null(A^T) are orthogonal complements.
- x=x_row(A)+x_null(A).
- b=b_col(A)+b_null(A^T).
- Ax=b is consistent iff b_null(A^T)=0.
- See Fig. 7.7.6.

Theorem 7.7.7 *Suppose that A is an $m \times n$ matrix and $\mathbf{b}$ is in the column space of A.*

(*a*) *If A has full column rank, then the system $A\mathbf{x} = \mathbf{b}$ has a unique solution, and that solution is in the row space of A.*

(*b*) *If A does not have full column rank, then the system $A\mathbf{x} = \mathbf{b}$ has infinitely many solutions, but there is a unique solution in the row space of A. Moreover, among all solutions of the system, the solution in the row space of A has the smallest norm.*

- Proof: (a) If A has a full rank, Theorem 7.5.6 implies Ax=b has a unique solution or is inconsistent. Since b is in col(A), it is consistent.
- (b). Theorem 7.5.6 implies Ax=0 has infinitely many solutions. Smallest norm -> omit.

Double perp theorem.

Theorem 7.7.8 (***Double Perp Theorem***) *If W is a subspace of R^n, then $(W^{\perp})^{\perp} = W$.*

- Proof: Show W is a subset of $(W^c)^c$:
  - Suppose w is in W. Then w is perp to every a in W^c. This means that w is in $(W^c)^c$.
- Show $(W^c)^c$ is a subset of W.
  - Let w be in $(W^c)^c$.
  - Write w=w_1+w_2, w_1 in W, w_2 in W^c.
  - w_2.w=0.
  - (w_2.w_1)+(w_2.w_2)=0. w_2.w_2=0 -> w_2 =0.

Orthogonal projection to W^c

- Proj_W^c(x) = x-proj_W(x) = Ix-Px=(I-P)x.
- Thus the matrix of Proj_W^c is $I-P=I-M(M^TM)^{-1}M^T$.
- I-P is also symmetric and idempotent.
- Rank(I-P)=n-rank P.