

# Orthogonality

- ♣ Motivation - More intuitive
- ♣ Inner product and projection
- ♣ Orthogonal vectors and linear independence
- ♣ Orthogonal complement
- ♣ Projection and least squares approximation
- ♣ Orthonormal bases and orthogonal matrices
- ♣ Gram-Schmidt orthogonalization process
- ♣ QR factorization

## Inner Product and Projection

**Definition:** An *inner product* on a real vector space  $V$  is a function that associates a real number  $\langle \mathbf{x}, \mathbf{y} \rangle$  with  $\mathbf{x}, \mathbf{y} \in V$  such that the following axioms are satisfied for all  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ ,  $c \in R$ .

- (1)  $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
- (2)  $\langle \mathbf{x} + \mathbf{z}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{z}, \mathbf{y} \rangle$
- (3)  $\langle c\mathbf{x}, \mathbf{y} \rangle = c\langle \mathbf{x}, \mathbf{y} \rangle$
- (4)  $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ ,  $\forall \mathbf{x}$  and  $\langle \mathbf{x}, \mathbf{x} \rangle = 0$  iff  $\mathbf{x} = \mathbf{0}$

◇ A real vector space with an inner product is called a real inner product space.

### Examples

- (1) In  $R^n$ , let  $\mathbf{x} = [x_1, x_2, \dots, x_n]^t$ ,  $\mathbf{y} = [y_1, y_2, \dots, y_n]^t$ , and define  $\langle \bullet, \bullet \rangle : R^n \times R^n \rightarrow R$  with  $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i = \mathbf{x}^t \mathbf{y}$ , then it is an inner product.
- (2) In  $C[0,1]$ , define  $\langle \bullet, \bullet \rangle : C[0,1] \times C[0,1] \rightarrow R$  with  $\langle f, g \rangle = \int_0^1 [f(x)g(x)]w(x)dx$ , then it is an inner product.

**Definition:** In a vector space  $V$ ,  $\mathbf{x}, \mathbf{y} \in V$  are orthogonal if  $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ .

**Definition:**  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  are said to be (mutually) orthogonal if  $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$  for  $i \neq j$ .

**Definition:**  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$  are (mutually) orthogonal and  $\|\mathbf{u}_i\|_2 = 1 \forall 1 \leq i \leq n$ , then  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$  are said to be orthonormal.

*Examples*

(1)  $[1, 0, 1]^t$ ,  $[1, 0, -1]^t$ , and  $[0, 1, 0]^t$  are orthogonal but not orthonormal in  $R^3$ .

(2) For  $f, g \in C[-1, 1]$ , define  $\langle f, g \rangle = \int_{-1}^1 [f(x)g(x)]dx$ , then

$\left\{ \frac{1}{\sqrt{2}}, \cos(\pi x), \sin(\pi x), \dots, \cos(n\pi x), \sin(n\pi x) \right\}$  is an orthonormal set.

(3) Let  $T_n \in C[-1, 1]$  be defined as  $T_n(x) = \cos(ncos^{-1}x)$ .  $\langle f, g \rangle = \int_{-1}^1 f(x)g(x) \frac{1}{\sqrt{1-x^2}} dx$  in  $C(0,1)$ . Then  $T_0 \equiv 1$ ,  $T_1(x) = x$ ,  $T_2(x) = 2x^2 - 1$ ,  $T_3(x) = 4x^3 - 3x$ , and in general,  $T_n(x)$  is polynomial of degree  $n$  with the leading coefficient  $2^{n-1}$ . Moreover,  $\{T_j's\}$  is an orthogonal set.

*Note that*

$$\begin{aligned} \langle T_m(x), T_n(x) \rangle &= \int_{-1}^1 \frac{T_m(x)T_n(x)}{\sqrt{1-x^2}} dx \\ &= \int_{-1}^1 \frac{\cos(m\cos^{-1}x) \cos(n\cos^{-1}x)}{\sqrt{1-x^2}} dx \\ &= - \int_{\pi}^0 \cos(my) \cos(ny) dy \quad \text{where } dy = \frac{1}{\sqrt{1-x^2}} dx \\ &= \int_0^{\pi} \cos(my) \cos(ny) dy \end{aligned}$$

Then  $\langle T_m(x), T_n(x) \rangle = 0$  if  $m \neq n$ ;  $\frac{\pi}{2}$  if  $m = n \geq 1$ ;  $\pi$  if  $m = n = 0$

## Orthogonal Vectors and Linear Independence

**Theorem:** If the nonzero vectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  are orthogonal, then they are linearly independent.

**Definition:** If  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  is a basis of  $R^n$  and they are orthonormal, then  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  is said to be an orthonormal basis of  $R^n$ .

**Definition:** Two subspaces  $V$  and  $W$  of  $R^n$  are orthogonal if  $\langle \mathbf{v}, \mathbf{w} \rangle = 0 \forall \mathbf{v} \in V$ , and  $\forall \mathbf{w} \in W$ , which can be written as  $V \perp W$ .

**Theorem:** The row space of a matrix  $A \in R^{m \times n}$  is orthogonal to its nullspace  $N(A)$ , i.e.,  $R(A^t) \perp N(A)$ . Similarly,  $R(A) \perp N(A^t)$ .

**Definition:** Given a subspace  $V$  of  $R^n$ , the space of all vectors orthogonal to  $V$  is called the *orthogonal complement* of  $V$ , and denoted by  $V^\perp$ .

### Fundamental Theorem of Linear Algebra

$$A \in R^{m \times n} \Rightarrow N(A) = R(A^t)^\perp \text{ and } N(A^t) = R(A)^\perp$$

### Direct Sum

Let  $V, W \in R^n$  such that  $W = V^\perp$ , then every  $\mathbf{x} \in R^n$  can be uniquely represented as  $\mathbf{x} = \mathbf{v} + \mathbf{w}$ , where  $\mathbf{v} \in V$  and  $\mathbf{w} \in W$ . Denote by  $R^n = V \oplus W$ .

## Intersection and Sum of Vector Spaces

**Definition:** The sum of vector subspaces  $V$  and  $W$  is defined as

$$V + W = \{\mathbf{v} + \mathbf{w} \mid \mathbf{v} \in V, \mathbf{w} \in W\}$$

**Definition:** Let  $V, W$  be subspaces of a vector space, then

(a)  $V \cup W$  is *in general* not a vector subspace

(b)  $V + W$  is a vector subspace

(c)  $V \cap W$  is a vector subspace

(d)  $V \cap V^\perp = \{\mathbf{0}\}$

(e)  $L \cap U = D$  and  $L + U = R^{n \times n}$ , where

$$L = \{A \in R^{n \times n} \mid a_{ij} = 0 \text{ if } i < j\}$$

$$U = \{B \in R^{n \times n} \mid a_{ij} = 0 \text{ if } i > j\}$$

$$D = \{C \in R^{n \times n} \mid a_{ij} = 0 \text{ if } i \neq j\}$$

**Theorem:**  $\dim(L + U) + \dim(L \cap U) = \dim(L) + \dim(U)$ .

## Projection Onto A Line

**Definition:**  $\mathbf{x}^t\mathbf{y} = \|\mathbf{x}\|_2\|\mathbf{y}\|_2\cos\theta$

**Definition:** The projection of a vector  $\mathbf{b}$  onto the line  $\mathbf{a}$  is a vector  $\gamma\mathbf{a}$  such that  $\langle \mathbf{b} - \gamma\mathbf{a}, \mathbf{a} \rangle = 0$ . Note that

$$\gamma\mathbf{a} = \frac{\mathbf{a}(\mathbf{a}^t\mathbf{b})}{\mathbf{a}^t\mathbf{a}} = \frac{\mathbf{a}\mathbf{a}^t}{\mathbf{a}^t\mathbf{a}}\mathbf{b} = P_a\mathbf{b}$$

where  $P_a$  is called the projection matrix along the line  $\mathbf{a}$ .

*Example:*  $\mathbf{a} = [1, 1, 1]^t$ ,  $\mathbf{b} = [1, 2, 3]^t$ , then  $P_a\mathbf{b} = 2\mathbf{a}$ .

Let  $P$  be a projection matrix along a certain line, then

- (a)  $P^2 = P$  and  $P^t = P$ .
- (b)  $P$  is not invertible ( $P$  has rank 1).
- (c) *Prove and Explain* why  $\mathbf{a}^t\mathbf{b} = 0$  implies that  $P_a + P_b = I$  and  $P_aP_b = O$ .
- (d) Show that  $\text{tr}(P) = 1$ .

## The Linear Least Squares Problems

Consider the problem of determining an  $\mathbf{x} \in R^n$  such that the *residual sum of squares*  $\rho^2(\mathbf{x}) = \|\mathbf{b} - A\mathbf{x}\|_2^2$  is minimized for given  $\mathbf{b} \in R^n$ ,  $A \in R^{n \times n}$ .

◇ *Best Line Fit:*

Given  $[x_i, y_i]^t \in R^2$  for  $1 \leq i \leq n$ , find a line which best fits these points. The problem is equivalent to finding  $m$  and  $b$  to minimize

$$f(m, b) = \sum_{i=1}^n (y_i - mx_i - b)^2$$

or to solve

$$\begin{bmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{bmatrix}$$

◇ *Best Parabola Fit:*

Given  $[x_i, y_i]^t \in R^2$  for  $1 \leq i \leq n$  ( $n = 7$ ), find a parabola which best fits these points. The problem is equivalent to finding  $a, b, c$  to minimize

$$f(a, b, c) = \sum_{i=1}^n (y_i - ax_i^2 - bx_i - c)^2$$

or to solve

$$\begin{bmatrix} \sum_{i=1}^n x_i^4 & \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^2 \\ \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i & \sum_{i=1}^n 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i^2 y_i \\ \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{bmatrix}$$

## Least Squares Approximations

**Theorem:** The linear least squares problem of minimizing  $\|\mathbf{b} - A\mathbf{x}\|_2$  always has a solution. The solution is unique iff  $\text{Null}(A) = \{\mathbf{0}\}$ .

*Corollary:* Let  $\mathbf{x}$  be a linear least squares solution of minimizing  $\|\mathbf{b} - A\mathbf{x}\|_2$ , then the residual vector  $\mathbf{r} = \mathbf{b} - A\mathbf{x}$  satisfies the *normal equations*.

$$A^t\mathbf{r} = A^t(\mathbf{b} - A\mathbf{x}) = \mathbf{0} \quad \text{or} \quad A^tA\mathbf{x} = A^t\mathbf{b}$$

**Theorem:**  $A\mathbf{x} = \mathbf{b}$  has a solution iff  $\mathbf{b} \in R(A)$ .

If the columns of  $A$  are linearly independent, then  $A^tA$  is invertible and  $\mathbf{x} = (A^tA)^{-1}A^t\mathbf{b}$ . The projection of  $\mathbf{b}$  onto the column space of matrix  $A$  is  $\mathbf{p} = A(A^tA)^{-1}A^t\mathbf{b}$ .

*Example:*

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \quad A^tA = \begin{bmatrix} 2 & 5 \\ 5 & 13 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \mathbf{p} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix}$$

**Theorem:** If  $A \in R^{m \times n}$  has rank  $n$  ( $n \leq m$ ), the normal equations  $A^tA\mathbf{x} = A^t\mathbf{b}$  has a unique solution  $\hat{\mathbf{x}} = (A^tA)^{-1}A^t\mathbf{b}$  and  $\hat{\mathbf{x}}$  is the unique LLS solution to  $A\mathbf{x} = \mathbf{b}$ .



## Orthonormal Basis and Orthogonal Matrices

**Definition:** The vectors  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$  are orthonormal if  $\|\mathbf{u}_k\|_2 = 1$ ,  $1 \leq k \leq n$ , and  $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0 \forall i \neq j$ .

**Definition:** An *orthogonal matrix* is simply a matrix with *orthonormal columns*. That is,  $Q \in R^{m \times k}$  is orthogonal if  $Q^t Q = I_k$ . In particular, if  $m = k$ , then  $Q^{-1} = Q^t$ .

*Examples:*

$$A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & -1 \end{bmatrix}, \quad P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{bmatrix}$$

- The column vectors of  $A$  and  $P$  are orthonormal
- The column vectors of  $B$  and  $H$  are orthogonal but not orthonormal

♣ *Some Properties of An Orthogonal Matrix  $Q \in R^{n \times n}$*

(a) *The columns of  $Q$  form an orthonormal basis for  $R^n$*

(b)  $Q^t Q = I$  and  $Q^{-1} = Q^t$

(c)  $\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2, \quad \forall \mathbf{x} \in R^n$

(d)  $\langle Q\mathbf{x}, Q\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle, \quad \forall \mathbf{x}, \mathbf{y} \in R^n$

(e)  $\|QA\|_2 = \|A\|_2, \quad \forall A \in R^{n \times k}$

(f)  $|Q| = 1$  or  $-1$

## Least Squares and Orthonormal Sets

**Theorem:** If the column vectors of  $A \in R^{m \times n}$  form an orthonormal set of vectors in  $R^m$ , then  $A^t A = I$  and the LLS solution to  $A\mathbf{x} = \mathbf{b}$  is  $\hat{\mathbf{x}} = (A^t A)^{-1} A^t \mathbf{b} = A^t \mathbf{b}$ .

**Theorem:** Let  $S$  be a subspace of an inner product vector space  $V$  and  $\mathbf{x} \in V$ . Let  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$  be an orthonormal basis for  $S$ . If

$$\mathbf{p} = \sum_{i=1}^n c_i \mathbf{u}_i, \quad \text{where } c_i = \langle \mathbf{x}, \mathbf{u}_i \rangle$$

Then,  $(\mathbf{x} - \mathbf{p}) \in S^\perp$

**Proof:**  $\langle \mathbf{x} - \mathbf{p}, \mathbf{u}_i \rangle = \langle \mathbf{x}, \mathbf{u}_i \rangle - \langle \mathbf{p}, \mathbf{u}_i \rangle = c_i - c_i = 0$

## Gram-Schmidt Orthogonalization Process

Let  $V = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$  be a set of independent vectors. The Gram-Schmidt process transforms the set  $V$  to an orthonormal set of  $U = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$  such that

$$\text{span}(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n) = \text{span}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$$

(a)  $\mathbf{q}_1 \leftarrow \mathbf{a}_1 / \|\mathbf{a}_1\|_2$

(b)  $\mathbf{t}_2 = \mathbf{a}_2 - \langle \mathbf{a}_2, \mathbf{q}_1 \rangle \mathbf{q}_1$ ;  $\mathbf{q}_2 \leftarrow \mathbf{t}_2 / \|\mathbf{t}_2\|_2$

(c)  $\mathbf{t}_k = \mathbf{a}_k - \sum_{i=1}^{k-1} \langle \mathbf{a}_k, \mathbf{q}_i \rangle \mathbf{q}_i$ ;  $\mathbf{q}_k \leftarrow \mathbf{t}_k / \|\mathbf{t}_k\|_2$  for  $3 \leq k \leq n$ .

Example:

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}; \quad \mathbf{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{q}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Example (QR Factorization):

$$A = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3] = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \sqrt{2} \\ 0 & \frac{1}{\sqrt{2}} & \sqrt{2} \\ 0 & 0 & 1 \end{bmatrix}$$

Example:

$$A = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3] = [\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3] \begin{bmatrix} \mathbf{q}_1^t \mathbf{a}_1 & \mathbf{q}_1^t \mathbf{a}_2 & \mathbf{q}_1^t \mathbf{a}_3 \\ 0 & \mathbf{q}_2^t \mathbf{a}_2 & \mathbf{q}_2^t \mathbf{a}_3 \\ 0 & 0 & \mathbf{q}_3^t \mathbf{a}_3 \end{bmatrix} = QR$$

## QR Factorization

**Theorem:** Every  $A \in R^{m \times n}$  with linearly independent columns can be factored into  $A = QR$ , where  $Q$  is orthogonal,  $R$  is upper- $\Delta$  and invertible.

**Proof:** Successively applied Householder matrices  $\{H'_j\}$  on  $A$ , we can get  $H_1 H_2 \cdots H_m A = R$ , where  $R$  is upper- $\Delta$ . If  $R$  is not invertible, then  $\exists \mathbf{x} \in R^n$  such that  $R\mathbf{x} = \mathbf{0}$ , then  $QR\mathbf{x} = \mathbf{0}$  and hence  $A\mathbf{x} = \mathbf{0}$  which contradicts that  $A$  has linearly independent column vectors.

Note: Suppose  $A = QR$ , the LLS solution of  $A\mathbf{x} = \mathbf{b}$  is reduced to solving a triangular system of equations  $R\mathbf{x} = Q^t\mathbf{b}$ .

Example:

$$A = \begin{bmatrix} 1 & -2 & -1 \\ 2 & 0 & 1 \\ 2 & -4 & 2 \\ 4 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0.2 & -0.4 & -0.8 \\ 0.4 & 0.2 & 0.4 \\ 0.4 & -0.8 & 0.4 \\ 0.8 & 0.4 & -0.2 \end{bmatrix} \begin{bmatrix} 5 & -2 & 1 \\ 0 & 4 & -1 \\ 0 & 0 & 2 \end{bmatrix} = QR$$

Let  $\mathbf{b} = [-1.4, 0.2, 1.2, -1.6]^t$ . By solving  $R\mathbf{x} = Q^t\mathbf{b}$ , we have  $\mathbf{x} = [-0.4, 0, 1]^t$  for the LLS solution of  $A\mathbf{x} = \mathbf{b}$ .

## Least Squares Approximation

(Matlab Code for Best Line Fitting)

```

% Script file: linefit.m
% A Linear Least Square Fit for (GPA_high school, GPA_university)
%
fin=fopen('dataGPA.txt','r');
fgetl(fin);
m=2; n=20;
T=fscanf(fin,'%f',[m n]);
fclose(fin);
T=T';
X=T(:,1);
Y=T(:,2);
A=[sum(X.*X), sum(X); sum(X), n];
b=[sum(X.*Y); sum(Y)];
v=A\b; % (0.8822, 0.0298)
for j=1:n,
    t=2.0+0.2*j;
    X1(j)=t;
    Y1(j)=t*v(1)+v(2);
end
plot(X1,Y1,'r-',X,Y,'bo'); axis([2 4 2 4]); grid;
legend('Best fitting line is y=0.8822x+0.0298','Location','NorthWest')
title('University GPA vs. High School GPA')
ylabel('University GPA')
xlabel('High School GPA')

```

### GPA Data Set *dataGPA.txt*

20 pairs of High School and University GPAs for Line Fit

```

3.75 3.19
3.45 3.34
2.87 2.23
3.60 3.46
3.42 2.97
4.00 3.79
2.65 2.55

```

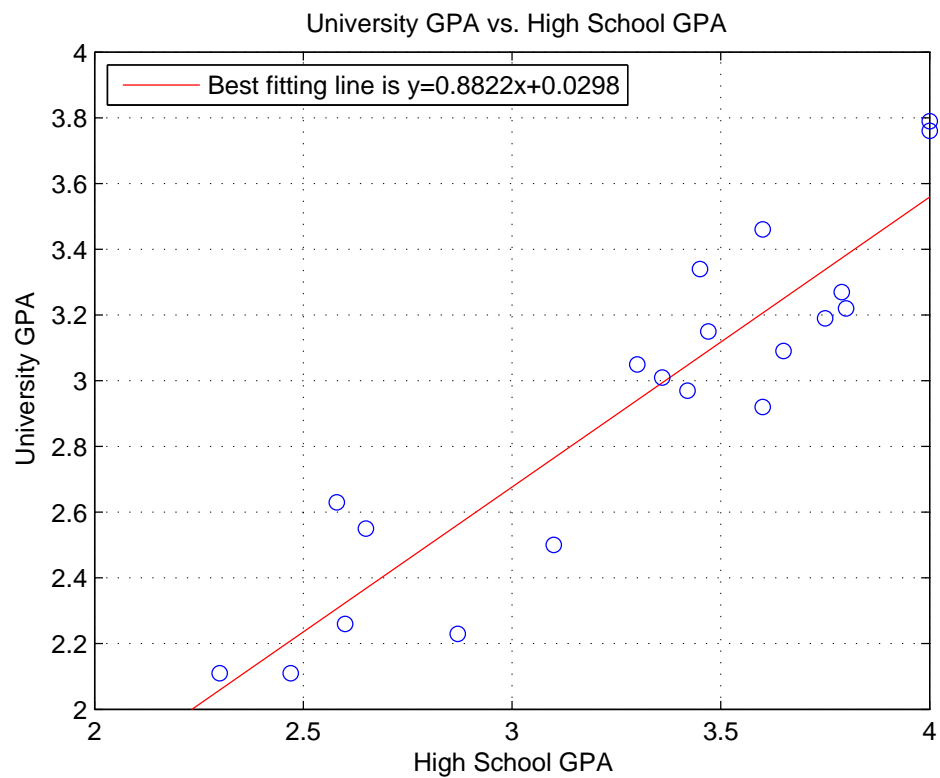


Figure 1: University GPA vs. High School GPA.

3.10 2.50  
 3.47 3.15  
 2.60 2.26  
 4.00 3.76  
 2.30 2.11  
 2.47 2.11  
 3.36 3.01  
 3.60 2.92  
 3.65 3.09  
 3.30 3.05  
 2.58 2.63  
 3.80 3.22  
 3.79 3.27

# Least Squares Approximation

## (Matlab Code for Best Parabola Fitting)

```

% Script file: plotQ.m
% A Quadratic Least Squares Fit for Data from Heath's book
%
fin=fopen('dataQua.txt','r');
fgetl(fin);
m=2; n=21;
T=fscanf(fin,'%f',[m n]);
fclose(fin);
T=T';
X=T(:,1);
Y=T(:,2);
[P S]=polyfit(X,Y,2);
ymu=mean(Y);
Yh=P(1)*X.^2 + P(2)*X + P(3);
top=norm(Yh-ymu,2);
bot=norm(Y-ymu,2);
R2=(top*top)/(bot*bot);
plot(X,Y,'ro',X,Yh,'b-');
legend('R^2 Statistics = 0.9335','\itY=-0.2384\itX^2+2.6704\itX+2.1757',4);
title('Quadratic Curve Fitting')

```

### Parabola Data Set *dataQua.txt*

Data for Parabola Fit (Quadratic Fit) from Heath's Book

0.0	2.9
0.5	2.7
1.0	4.8
1.5	5.3
2.0	7.1
2.5	7.6
3.0	7.7
3.5	7.6
4.0	9.4
4.5	9.0
5.0	9.6

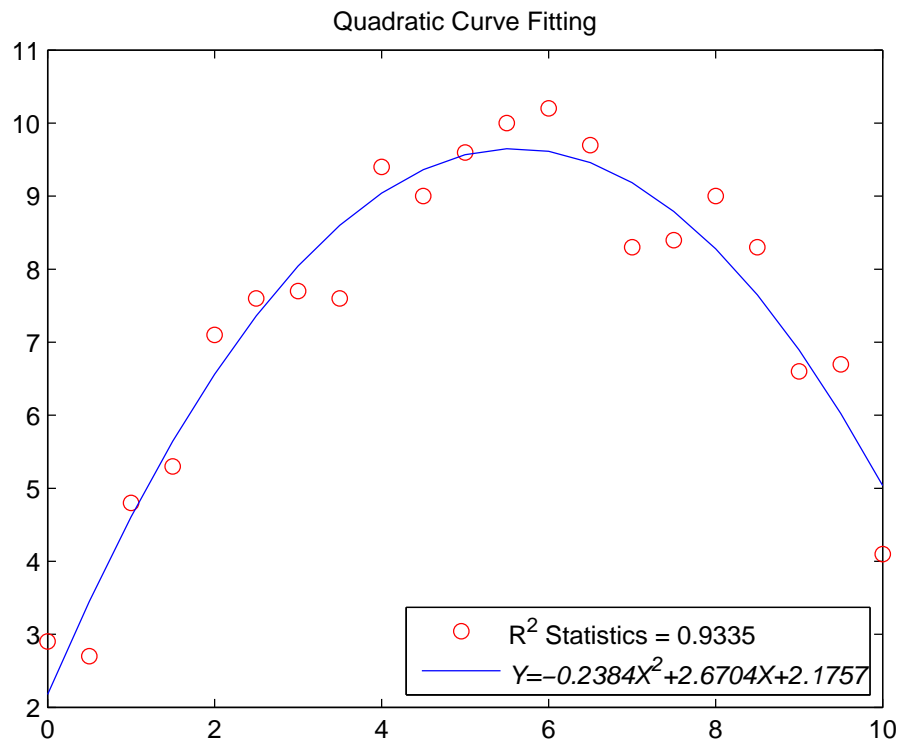


Figure 2: Best Parabola Fitting.

5.5 10.0  
6.0 10.2  
6.5 9.7  
7.0 8.3  
7.5 8.4  
8.0 9.0  
8.5 8.3  
9.0 6.6  
9.5 6.7  
10.0 4.1