# §6.6 Applications to Linear Models

## * Least-squares Lines

least-squares line is the line $y = \beta_0 + \beta_1 x$ that minimize $\|X\vec{\beta} - \vec{y}\|$.

where

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \cdots & \\ 1 & x_n \end{pmatrix}, \quad \vec{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

It is equivalent to find the least-squares solution for equation $X\vec{\beta} = \vec{y}$.

## * The General Linear Model

Fit data by curves that have the general form

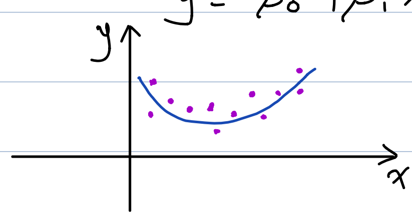
$$y = \beta_0 f_0(x) + \beta_1 f_1(x) + \cdots + \beta_k f_k(x)$$

where $f_0, \cdots, f_k$ are known functions and $\beta_0, \cdots, \beta_k$ are parameters that must be determined.

Example: In ecology, a parabolic curve that opens downward is used to model the net primary production of nutrients in a plant. Given $(x_1, y_1), \cdots, (x_n, y_n)$, describe the linear model that produces a "least-squares fit" of the data by equation

$$y = \beta_0 + \beta_1 x + \beta_2 x^2.$$

Solution:



Suppose the actual values of the parameters are $\beta_0, \beta_1, \beta_2$. Then the coordinates of the first data point $(x_1, y_1)$ satisfy an equation of the form

$$y_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \varepsilon_1$$

where $\varepsilon_1$ is the residual error between the observed value $y_1$ and the predicted $y$-value $\beta_0 + \beta_1 x_1 + \beta_2 x_1^2$.

Similarly, we can get the following equations

$$y_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \varepsilon_1$$
$$y_2 = \beta_0 + \beta_1 x_2 + \beta_2 x_2^2 + \varepsilon_2$$
$$y_3 = \beta_0 + \beta_1 x_3 + \beta_2 x_3^2 + \varepsilon_3$$
$$\vdots$$
$$y_n = \beta_0 + \beta_1 x_n + \beta_2 x_n^2 + \varepsilon_n$$

In matrix form, we have

$$\vec{y} = X\vec{\beta} + \vec{\varepsilon}$$

where $\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, $X = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ & \cdots & \\ 1 & x_n & x_n^2 \end{pmatrix}$, $\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$

Exercise: Given the data $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ that comes from a company's total costs. The equation of the form

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$$

is an appropriate model to fit the data. Describe the linear model that gives a least-squares fit of the data.

Solution: $\vec{y} = X\beta + \varepsilon$, where

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & x_n^3 \end{pmatrix}, \quad \vec{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}, \quad \vec{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Example: In geography, local models of terrain are constructed from data $(u_1, v_1, y_1), \cdots, (u_n, v_n, y_n)$, where $u_j$, $v_j$ and $y_j$ are latitude, longitude, and altitude, respectively. Describe the linear model based on

$$y = \beta_0 + \beta_1 u + \beta_2 v$$

that gives a least-squares fit to such data. The solution is called the least-squares plane.

Solution: We expect the data to satisfy the following equations:

$$y_1 = \beta_0 + \beta_1 u_1 + \beta_2 v_1 + \varepsilon_1$$
$$y_2 = \beta_0 + \beta_1 u_2 + \beta_2 v_2 + \varepsilon_2$$
$$\vdots$$
$$y_n = \beta_0 + \beta_1 u_n + \beta_2 v_n + \varepsilon_n$$

The matrix form is

$$\vec{y} = X\vec{\beta}, \quad \text{where}$$

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & u_1 & v_1 \\ 1 & u_2 & v_2 \\ \vdots & \vdots & \vdots \\ 1 & u_n & v_n \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}, \quad \vec{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

## §7.1 Diagonalization of Symmetric Matrices

Def: A symmetric matrix is a matrix $A$ such that $A^T = A$.

Example:

Symmetric $\begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix}$, $\begin{pmatrix} 0 & -1 & 0 \\ -1 & 5 & 8 \\ 0 & 8 & -7 \end{pmatrix}$, $\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$

Nonsymmetric: $\begin{pmatrix} 1 & -3 \\ 3 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & -4 & 0 \\ -6 & 1 & -4 \\ 0 & -6 & 1 \end{pmatrix}$, $\begin{pmatrix} 5 & 4 & 3 & 2 \\ 4 & 3 & 2 & 1 \\ 3 & 2 & 1 & 0 \end{pmatrix}$

Example: If possible, diagonalize the matrix

$$A = \begin{pmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{pmatrix}$$

Solution: $\det(A - \lambda I) = \begin{vmatrix} 6-\lambda & -2 & -1 \\ -2 & 6-\lambda & -1 \\ -1 & -1 & 5-\lambda \end{vmatrix}$

$$= - \begin{vmatrix} -1 & -1 & 5-\lambda \\ -2 & 6-\lambda & -1 \\ 6-\lambda & -2 & -1 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 1 & \lambda-5 \\ -2 & 6-\lambda & -1 \\ 6-\lambda & -2 & -1 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 1 & \lambda-5 \\ 0 & 8-\lambda & 2\lambda-11 \\ 0 & \lambda-8 & (\lambda-5)(\lambda-6)-1 \end{vmatrix}$$

$$= \begin{vmatrix} 8-\lambda & 2\lambda-11 \\ \lambda-8 & \lambda^2-11\lambda+29 \end{vmatrix}$$

$$= (8-\lambda) \begin{vmatrix} 1 & 2\lambda-11 \\ -1 & \lambda^2-11\lambda+29 \end{vmatrix}$$

$$= (8-\lambda)\,[\lambda^2-11\lambda+29+2\lambda-11]$$

$$= (8-\lambda)\,[\lambda^2-9\lambda+18]$$

$$= (8-\lambda)(\lambda-3)(\lambda-6)$$

Hence the eigenvalues are $\lambda_1 = 8$, $\lambda_2 = 3$, $\lambda_3 = 6$

For $\lambda_1 = 8$, solve $\begin{pmatrix} 6-8 & -2 & -1 \\ -2 & 6-8 & -1 \\ -1 & -1 & 5-8 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

$$\begin{pmatrix} -2 & -2 & -1 \\ -2 & -2 & -1 \\ -1 & -1 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -2 & -2 & -1 & 0 \\ -2 & -2 & -1 & 0 \\ -1 & -1 & -3 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 3 & 0 \\ 2 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 1 & 3 & 0 \\ 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \begin{cases} x_1 + x_2 + 3x_3 = 0 \\ \qquad\quad -5x_3 = 0 \end{cases}$$

$x_1 = -x_2$

$x_3 = 0$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_2 \\ 0 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

Let $\vec{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, $\vec{u}_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$

For $\lambda_2 = 6$, we can get $\vec{v}_2 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$, $\vec{u}_2 = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix}$

For $\lambda_3 = 3$, $\vec{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\vec{u}_3 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$

Let $P = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}$, $D = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

Then $A = PDP^{-1}$

Remark: In this example, $P^TP = I$. Actually

$$P^TP = \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

**Thm:** If $A$ is symmetric, then any two eigenvectors from different eigenspaces are orthogonal.

**Proof:** Let $\vec{v}_1, \vec{v}_2$ be eigenvectors that correspond to distinct eigenvalues, say, $\lambda_1$ and $\lambda_2$. Then

$$\begin{aligned}
\lambda_1 \vec{v}_1 \cdot \vec{v}_2 &= A\vec{v}_1 \cdot \vec{v}_2 \\
&= (A\vec{v}_1)^T \vec{v}_2 \\
&= \vec{v}_1^{\,T} A^T \vec{v}_2 \\
&= \vec{v}_1^{\,T} A \vec{v}_2 \\
&= \vec{v}_1^{\,T} \lambda_2 \vec{v}_2 \\
&= \lambda_2 \vec{v}_1 \cdot \vec{v}_2
\end{aligned}$$

Hence $(\lambda_1 - \lambda_2)\vec{v}_1 \cdot \vec{v}_2 = 0$.

But $\lambda_1 - \lambda_2 \neq 0$, so $\vec{v}_1 \cdot \vec{v}_2 = 0$.

**Def:** An $n \times n$ matrix $A$ is said to be orthogonally diagonalizable if there are an orthogonal matrix $P$ (with $P^{-1} = P^T$) and a diagonal matrix $D$ such that

$$A = PDP^{-1} = PDP^T$$

**Remark:** If $A$ is orthogonally diagonalizable, then

$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$$

that is, $A$ is symmetric.

**Thm:** An $n \times n$ matrix $A$ is orthogonally diagonalizable if and only if $A$ is a symmetric matrix.

Example: Orthogonally diagonalize the matrix

$$A = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{pmatrix},$$

whose characteristic equation is

$$0 = -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda-7)^2(\lambda+2).$$

Solution: The usual calculations produce bases for the eigenspaces:

$\lambda = 7$: $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}$;

$\lambda = -2$, $\vec{v}_3 = \begin{pmatrix} -1 \\ -\frac{1}{2} \\ 1 \end{pmatrix}$

$$\vec{u}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\vec{v}_2{}' = \vec{v}_2 - (\vec{v}_2 \cdot \vec{u}_1) \cdot \vec{u}_1$$

$$= \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix} - \left(-\frac{1}{2} \cdot \frac{1}{\sqrt{2}} + 0 + 0\right) \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{1}{4} \\ 0 \\ \frac{1}{4} \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{1}{4} \\ 1 \\ \frac{1}{4} \end{pmatrix}$$

$$\vec{u}_2 = \frac{\vec{v}_2{}'}{\|\vec{v}_2{}'\|} = \begin{pmatrix} -\frac{1}{\sqrt{18}} \\ \frac{4}{\sqrt{18}} \\ \frac{1}{\sqrt{18}} \end{pmatrix}$$

$$\vec{u}_3 = \frac{\vec{v}_3}{\|\vec{v}_3\|} = \begin{pmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{pmatrix}$$

$$P = [\vec{u}_1, \vec{u}_2, \vec{u}_3] = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{18}} & -\frac{2}{3} \\ 0 & \frac{4}{\sqrt{18}} & -\frac{1}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{18}} & \frac{2}{3} \end{pmatrix}$$

$$D = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Then $P$ orthogonally diagonalizes $A$, and $A = PDP^T$.

**Thm: The Spectral Theorem for Symmetric Matrices**

An $n \times n$ symmetric matrix $A$ has the following properties:

a) $A$ has $n$ real eigenvalues, counting multiplicities

b) The dimension of the eigenspace for each eigenvalue $\lambda$ equals the multiplicity of $\lambda$ as a root of the characteristic equation

c) The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal.

d) A is orthogonally diagonalizable.

## * Spectral Decomposition

When A is symmetric, it is orthoganally diagonalizable.

$$A = PDP^T$$

$$= [\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_n] \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} \vec{u}_1 \\ \vec{u}_2 \\ \vdots \\ \vec{u}_n \end{pmatrix}$$

$$= [\lambda_1\vec{u}_1, \lambda_2\vec{u}_2, \cdots, \lambda_n\vec{u}_n] \begin{pmatrix} \vec{u}_1 \\ \vec{u}_2 \\ \vdots \\ \vec{u}_n \end{pmatrix}$$

$$= \lambda_1 \vec{u}_1\vec{u}_1^T + \lambda_2 \vec{u}_2\vec{u}_2^T + \cdots + \lambda_n \vec{u}_n\vec{u}_n^T$$

Def: $A = \lambda_1 \vec{u}_1\vec{u}_1^T + \lambda_2 \vec{u}_2\vec{u}_2^T + \cdots + \lambda_n \vec{u}_n\vec{u}_n^T$ is called a spectral decomposition of A.

Ex: Show that if A is orthogonally diagonalizable, then so is $A^2$.

Solution: If A is orthogonally diagonalizable, then A is symmetric. That is, $A^T = A$.

Hence $(A^2)^T = (A \cdot A)^T = A^T \cdot A^T = A \cdot A = A^2$

$A^2$ is symmetric. Therefore, $A^2$ is orthogonally diagonalizable.