1.6 and 5.3. Curve Fitting

One of the broadest applications of linear algebra is to curve fitting, especially in determining unknown coefficients in functions. You should know that, given two points in the plane, there is a line passing through both, and you should also be able to find an equation for that line.

To make things slightly more complicated, given three points in the plane, none of which is directly over another, there is a (possibly degenerate) parabola passing through those three points. Finding an equation for that parabola involves solving for the coefficients in the equation $y = ax^2 + bx + c$.

Example. Find an equation for the parabola passing through the points $(1, 2)$, $(2, 4)$, and $(3, 8)$.

Example. Find an equation for the parabola passing through the points $(1, 2)$, $(2, 4)$, and $(3, 8)$.

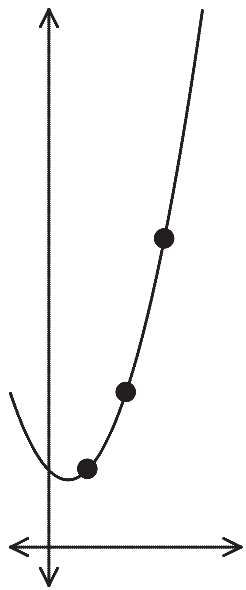
Using the equation $y = ax^2 + bx + c$, we require that

$$\begin{aligned} a &+ b + c = 2 \\ 4a &+ 2b + c = 4 \\ 9a &+ 3b + c = 8 \end{aligned}$$

Since the "coefficient matrix" is invertible, we have

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}.$$

So the parabola is $y = ax^2 + bx + c = x^2 - x + 2$**.**



We could also find a linear combination of $\frac{1}{x}$**,** $\cos(\pi x)$**, and** x **that passes through these three points as well.**

Example. Find an equation of the form

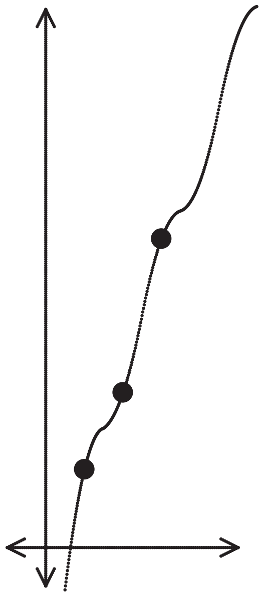
$$y = \frac{a}{x} + b\cos(\pi x) + cx$$

whose graph passes through the points $(1,2)$**,** $(2,4)$**, and** $(3,8)$**.**

We obtain the new system of linear equations:

$$\begin{aligned} a - b + c &= 2 \\ \frac{1}{2}a + b + 2c &= 4 \\ \frac{1}{3}a - b + 3c &= 8 \end{aligned}$$

Then $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 1/2 & 1 & 2 \\ 1/3 & -1 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix} = \begin{bmatrix} -6/5 \\ -3/5 \\ 13/5 \end{bmatrix}$**, so the desired curve is** $y = \dfrac{-6}{5x} - \dfrac{3}{5}\cos(\pi x) + \dfrac{13}{5}\,x$**.**



Of course, no line $y = ax + b$ is guaranteed to pass through three given points in the plane (such as $(1,2)$, $(2,4)$, and $(3,8)$). However, we can ask for the line which comes closest to those three points, where "close" is defined as

$$((ax_1 + b) - y_1)^2 + ((ax_2 + b) - y_2)^2 + \cdots + ((ax_k + b) - y_k)^2,$$

where the points are (x_1, y_1), $\ldots$, (x_k, y_k).

The line we are looking for is the least squares solution to the system of linear equations whose individual equations are

$$x_i a + b = y_i.$$

(Remember, a and b are the "variables" here.)

For our three points $(1,2)$, $(2,4)$, and $(3,8)$, the system is

$$\begin{aligned} a + b &= 2 \\ 2a + b &= 4 \\ 3a + b &= 8 \end{aligned}$$

so that $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}$ **and** $B = \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix}$**. Note that the first column of** A **is the** x **values, the second column is all 1's, and the vector** B **is the** y **values. This is because the equation we're trying to fit is** $(x)a + (1)b = (y)$**.**

Then we need to solve the equation $(A^\top A)\begin{bmatrix} a \\ b \end{bmatrix} = (A^\top B)$ **for** a **and** b**.**

$$A^\top A = \begin{bmatrix} 14 & 6 \\ 6 & 3 \end{bmatrix}$$

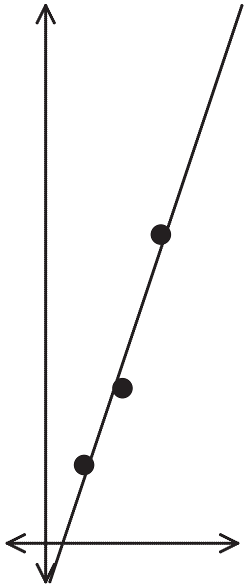
$$A^\top B = \begin{bmatrix} 34 \\ 14 \end{bmatrix}$$

$$(A^\top A)^{-1} = \frac{1}{6} \begin{bmatrix} 3 & -6 \\ -6 & 14 \end{bmatrix}$$

$$\begin{bmatrix} a \\ b \end{bmatrix} = (A^\top A)^{-1}(A^\top B) = \begin{bmatrix} 3 \\ -4/3 \end{bmatrix}$$

So the line is $y = 3x - \frac{4}{3}$. How did we do?

The graph of $y = 3x - \frac{4}{3}$ along with the data points:



Finding the best-fitting line is also called Linear Regression.

For the data points $(8, 10)$**,** $(7, 9)$**,** $(7, 5)$**,** $(4, 4)$**, and** $(1, 3)$**, find the best-fitting:**

- **Line** $y = ax + b$
- **Parabola** $y = ax^2 + bx + c$
- **Parabola without a constant term** $y = ax^2 + bx$
- **Parabola without a linear term** $y = ax^2 + c$
- **Weird-type curve from Section 1.6:**

$$y = \frac{a}{x} + b\cos(\pi x) + cx$$

- **Circle** $(x - h)^2 + (y - k)^2 = r^2$ **(Yes, this can be done!!!)**
- **Exponential curve** $y = C \cdot a^x$
- **Power function curve** $y = a \cdot x^b$

For the line $(y) = ax + b = (x)a + (1)b$: The system to find least squares for is $AX = B$, where

$$A = \begin{bmatrix} 8 & 1 \\ 7 & 1 \\ 7 & 1 \\ 4 & 1 \\ 1 & 1 \end{bmatrix} \quad \textbf{and} \quad B = \begin{bmatrix} 10 \\ 9 \\ 5 \\ 4 \\ 3 \end{bmatrix}.$$

$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0.8916 \\ 1.3855 \end{bmatrix}$**, so** $y = 0.8916x + 1.3855$**.**

For the parabola $(y) = ax^2 + bx + c = (x^2)a + (x)b + (1)c$**:**
The system to find least squares for is $AX = B$**, where**

$$A = \begin{bmatrix} 64 & 8 & 1 \\ 49 & 7 & 1 \\ 49 & 7 & 1 \\ 16 & 4 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \textbf{and} \quad B = \begin{bmatrix} 10 \\ 9 \\ 5 \\ 4 \\ 3 \end{bmatrix}.$$

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0.1882 \\ -0.7888 \\ 3.7213 \end{bmatrix}, \textbf{ so } y = 0.1882x^2 - 0.7888x + 3.7213.$$

For the parabola with no constant term ($(y) = ax^2 + bx = (x^2)a + (x)b$): The system to find least squares for is $AX = B$, where

$$A = \begin{bmatrix} 64 & 8 \\ 49 & 7 \\ 49 & 7 \\ 16 & 4 \\ 1 & 1 \end{bmatrix} \quad \textbf{and} \quad B = \begin{bmatrix} 10 \\ 9 \\ 5 \\ 4 \\ 3 \end{bmatrix}.$$

$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0.0123 \\ 1.0138 \end{bmatrix}$, **so** $y = 0.0123x^2 + 1.0138x$**.**

For the parabola with no linear term $((y) = ax^2 + c = (x^2)a + (1)c)$: The system to find least squares for is $AX = B$, where

$$A = \begin{bmatrix} 64 & 1 \\ 49 & 1 \\ 49 & 1 \\ 16 & 1 \\ 1 & 1 \end{bmatrix} \quad \textbf{and} \quad B = \begin{bmatrix} 10 \\ 9 \\ 5 \\ 4 \\ 3 \end{bmatrix}.$$

$$\begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} 0.1031 \\ 2.5090 \end{bmatrix}, \textbf{ so } y = 0.1031x^2 + 2.5090.$$

For the Weird-type curve from Section 1.6: $y = \dfrac{a}{x} + b\cos(\pi x) + cx$**: The system to find least squares for is** $AX = B$**, where**

$$A = \begin{bmatrix} 1/8 & 1 & 8 \\ 1/7 & -1 & 7 \\ 1/7 & -1 & 7 \\ 1/4 & 1 & 4 \\ 1 & -1 & 1 \end{bmatrix} \quad \textbf{and} \quad B = \begin{bmatrix} 10 \\ 9 \\ 5 \\ 4 \\ 3 \end{bmatrix}.$$

$\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ 🖩 $= \begin{bmatrix} 2.1371 \\ 0.4180 \\ 1.0479 \end{bmatrix}$**, so** $y = \dfrac{2.1371}{x} + 0.4180\cos(\pi x) + 1.0479x$**.**

For the circle, we have to do something clever. The equation of the circle with center (h, k) and radius r is

$$(x-h)^2+(y-k)^2=r^2,$$

which can be rewritten as

$$(2x)h+(2y)k+(r^2-h^2-k^2)=(x^2+y^2)$$

which is not linear in h, k, and r. However — and this is the trick — the equation IS linear with respect to h, k, and $L=r^2-h^2-k^2$. So what we do is to find the least squares solution to the linear equation

$$(2x)h+(2y)k+(1)L=(x^2+y^2)$$

and then get r from h, k, and L.

The circle's system of equations becomes $AX = B$, where

$$A = \begin{bmatrix} 16 & 20 & 1 \\ 14 & 18 & 1 \\ 14 & 10 & 1 \\ 8 & 8 & 1 \\ 2 & 6 & 1 \end{bmatrix} \quad \textbf{and} \quad B = \begin{bmatrix} 8^2 + 10^2 \\ 7^2 + 9^2 \\ 7^2 + 5^2 \\ 4^2 + 4^2 \\ 1^2 + 3^2 \end{bmatrix} = \begin{bmatrix} 164 \\ 130 \\ 74 \\ 32 \\ 10 \end{bmatrix}.$$

Then $\begin{bmatrix} h \\ k \\ L \end{bmatrix} = \begin{bmatrix} 2.6718 \\ 8.2194 \\ -48.7767 \end{bmatrix}$**, so** $r = \sqrt{h^2 + k^2 + L} \approx 5.0913$**.**

$$(x - 2.6718)^2 + (y - 8.2194)^2 = 5.0913^2$$

For the exponential curve $y = C \cdot a^x$, take the natural log of both sides, to get an equation linear in $\ln C$ and $\ln a$:

$$(\ln y) = \ln(C \cdot a^x) = \ln C + x \ln a = (1)L + (x)M$$

$$A = \begin{bmatrix} 1 & 8 \\ 1 & 7 \\ 1 & 7 \\ 1 & 4 \\ 1 & 1 \end{bmatrix} \quad \textbf{and} \quad B = \begin{bmatrix} \ln 10 \\ \ln 9 \\ \ln 5 \\ \ln 4 \\ \ln 3 \end{bmatrix} = \begin{bmatrix} 2.3026 \\ 2.1972 \\ 1.6094 \\ 1.3862 \\ 1.0986 \end{bmatrix}$$

Then $\begin{bmatrix} L \\ M \end{bmatrix} = \begin{bmatrix} 0.8563 \\ 0.1597 \end{bmatrix}$**, so** $C = e^{0.8563} \approx 2.3545$**,** $a = e^{0.1597} \approx 1.1732$**, and** $y = 2.3545 \cdot 1.1732^x$**.**

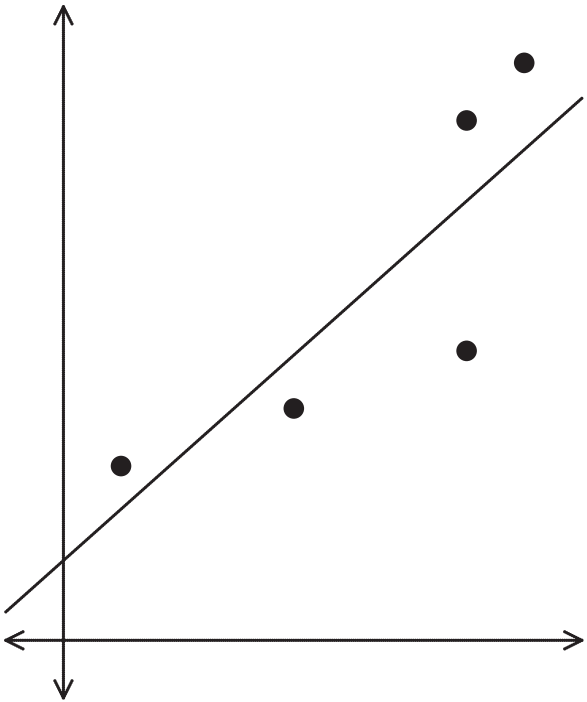
For the power function curve $y = a \cdot x^b$, take the natural log of both sides, to get an equation linear in $\ln a$ and b:

$$(\ln y) = \ln(a \cdot x^b) = \ln a + b \ln x = (1)L + (\ln x)b$$

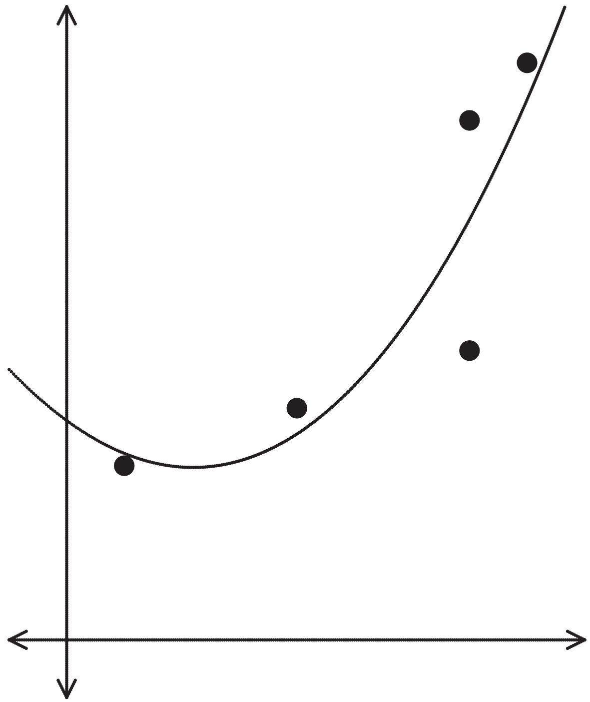
$$A = \begin{bmatrix} 1 & \ln 8 \\ 1 & \ln 7 \\ 1 & \ln 7 \\ 1 & \ln 4 \\ 1 & \ln 1 \end{bmatrix} \quad \textbf{and} \quad B = \begin{bmatrix} \ln 10 \\ \ln 9 \\ \ln 5 \\ \ln 4 \\ \ln 3 \end{bmatrix} = \begin{bmatrix} 2.3026 \\ 2.1972 \\ 1.6094 \\ 1.3862 \\ 1.0986 \end{bmatrix}$$

Then $\begin{bmatrix} L \\ b \end{bmatrix} = \begin{bmatrix} 0.9955 \\ 0.4916 \end{bmatrix}$**, so** $a = e^{0.9955} \approx 2.706$**,**
and $y = 2.706 \cdot x^{0.4916}$**.**

And now for the pictures:

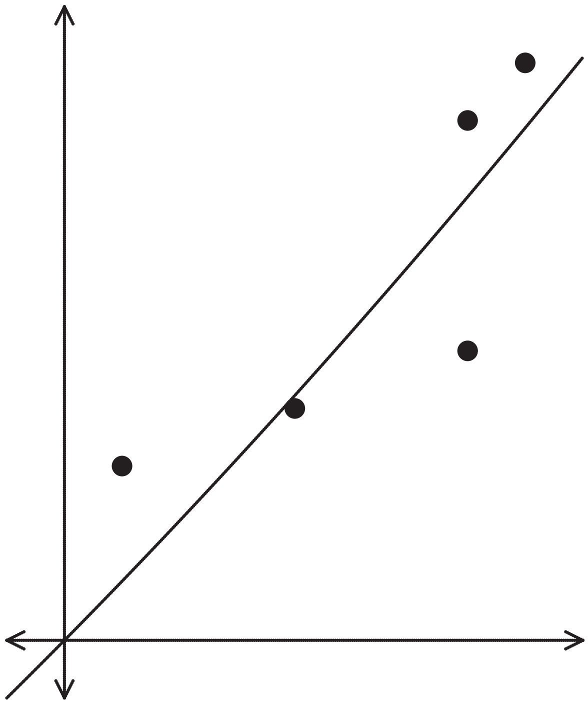


$y = 0.8916x + 1.3855$

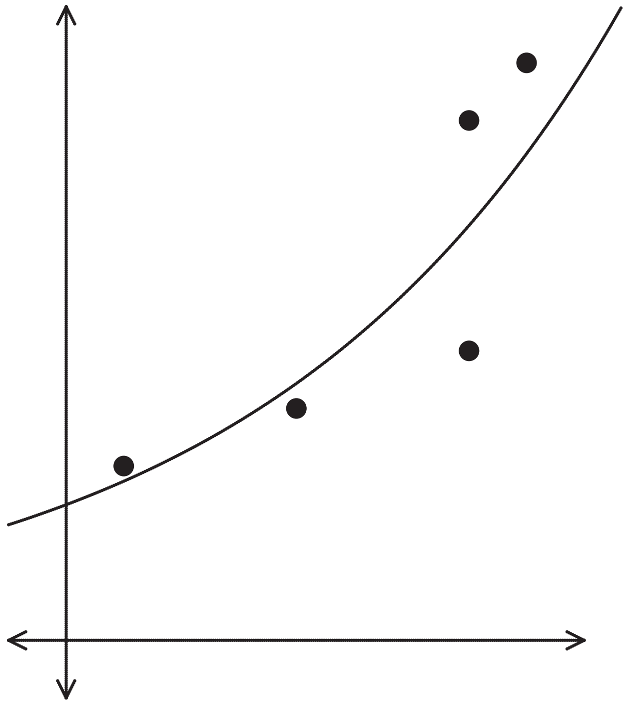


$y = 0.1882x^2 - 0.7888x + 3.7213$

More pictures:

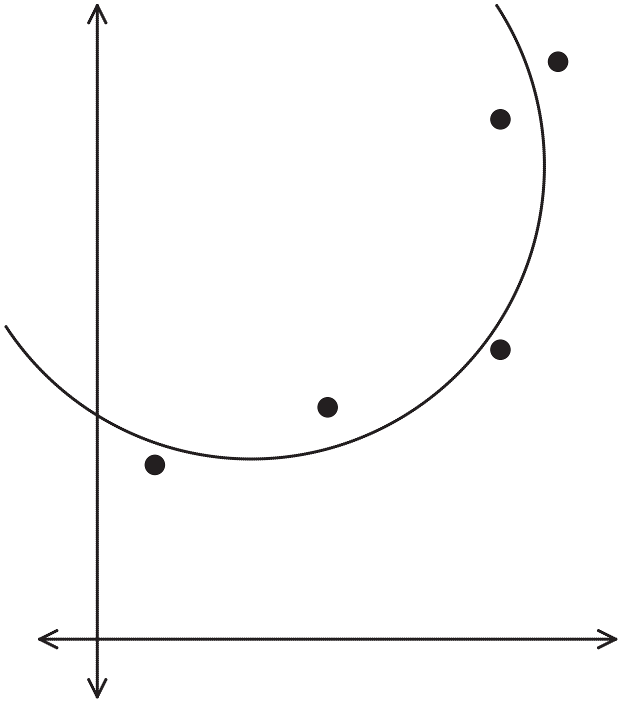


$y = 0.0123x^2 + 1.0138x$

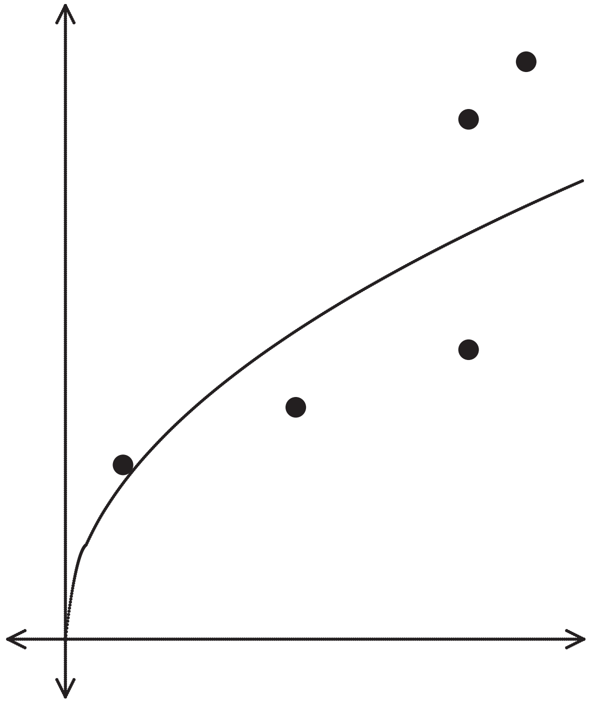


$y = 2.3545 \cdot 1.1732^x$

Even more pictures:

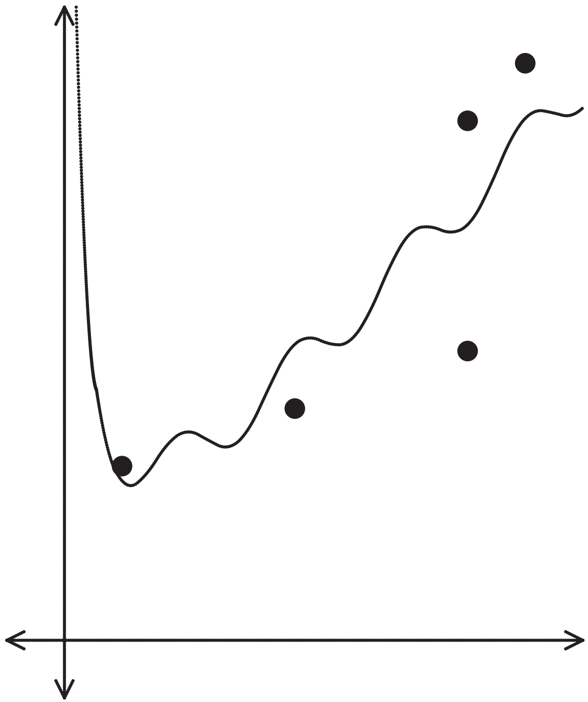


$$(x - 2.6718)^2 + (y - 8.2194)^2 = 5.0913^2$$

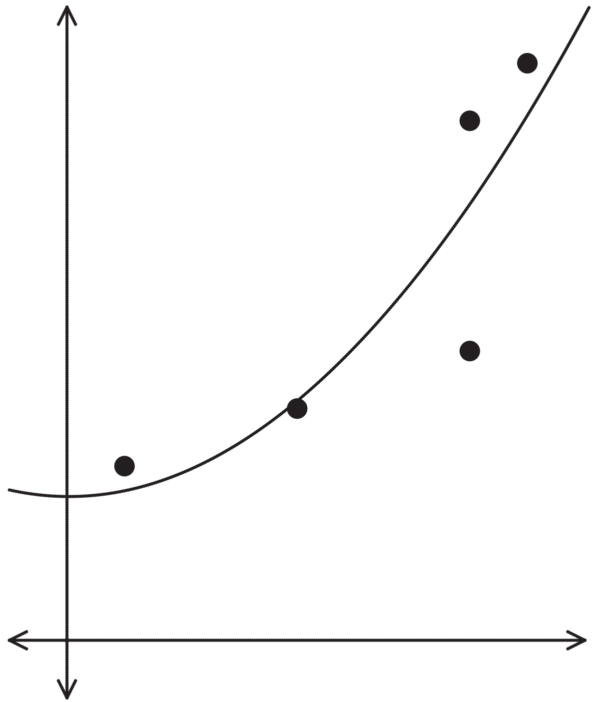


$$y = 2.706 \cdot x^{0.4916}$$

Final pictures:



$$y = \frac{2.1371}{x} + 0.4180\cos(\pi x) + 1.0479x$$



$$y = 0.1031x^2 + 2.5090$$

For Test 3, you should know how to:

- **find coordinates of a vector with respect to an orthogonal basis** ($c_i = \dfrac{u \cdot v_i}{v_i \cdot v_i}$) **[5.1]**
- **find a basis for the orthogonal complement of a subspace** W**, given a basis for** W **[5.1]**
- **find the vector closest to a subspace** W**, given a basis for** W **(perhaps orthogonal, perhaps not) [5.1, 5.2]**
- **find the Least Squares Solution to a system of linear equations [5.2]**
- **find the line, or parabola, or a type of parabola, that fits or best fits a list of data points [1.6, 5.3]**

$\Longrightarrow$

For Test 3, you should know how to:

- **find an orthogonal basis for a subspace W, given a "generic" basis (Gram-Schmidt) [5.4]**
- **find an orthonormal basis for a subspace W, given an orthogonal basis [5.4]**

□