CS 131 Computer Vision: Foundations and Applications (Fall 2016)
Model Fitting Using Least Squares

An extremely powerful technique in computer vision is *least squares model fitting*. This allows us to use data to choose the parameters of a model. As a concrete example, suppose that we are given a set of data points (x_i, y_i) and we believe that these points approximately lie on some parabola which could be described by the equation

$$y = ax^2 + bx + c$$

In order to *fit* this model to our data we need to determine the values of the parameters a, b, and c. If each of our data points were perfectly described by the model then the following system of equations would hold:

$$\begin{aligned} ax_1^2 + bx_1 + c &= y_1 \\ ax_2^2 + bx_2 + c &= y_2 \\ ax_3^2 + bx_3 + c &= y_3 \\ &\vdots \end{aligned}$$

We can rewrite this system of equations in matrix form:

$$\begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \end{bmatrix}$$

If we let

$$A = \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \\ \vdots & \vdots & \vdots \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \end{bmatrix}$$

Then this system of equations can simply be written as

$$A\mathbf{x} = \mathbf{y}$$

In this system of equations it is important to note that the values $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ are the coordinates of the data points that we were given, so these are known values. The only unknown variables in the above system of equations are the model parameters a, b, and c, which are contained in the vector $\mathbf{x}$.

If the matrix A were square and invertible, then we could simply solve for the model parameters by choosing

$$\mathbf{x} = A^{-1}\mathbf{y}$$

However, in most cases the matrix A will be neither square nor invertible. When this happens then the equation $A\mathbf{x} = \mathbf{y}$ has no solution. In this case we want to find the model parameters $\mathbf{x}$ so that $A\mathbf{x} \approx \mathbf{y}$ with as little error as possible. One way of choosing x in this situation is to choose the value of $\mathbf{x}$ that minimizes the *least squares error* defined by

$$error = \|A\mathbf{x} - \mathbf{y}\|^2$$

You do not need to know how to choose the value of $\mathbf{x}$ that minimizes this error. Instead, you should know that the \ operator in MATLAB will solve this problem for you. If the MATLAB variables `A` and `y` contain the data from A and $\mathbf{y}$ then you can solve for the best value of $\mathbf{x}$ with the following MATLAB command:

```
x = A \ y
```

In summary, we can find the parabola that best describes our data points (x_i, y_i) as follows:

1. Form the matrix A whose ith row is $(x_i^2, x_1, 1)$.
2. Form the vector $\mathbf{y}$ whose ith element is y_i.
3. Solve for the model parameters $\mathbf{x}$ using MATLAB: `x = A \ y`
4. Unpack the vector $\mathbf{x}$ into the individual model parameters (a, b, c) using the fact that $\mathbf{x} = (a, b, c)$.
5. Our final parabola that best fits the data is given by $y = ax^2 + bx + c$.