02-512 Assignment 02

Karan Sikka
ksikka@cmu.edu
October 1, 2014

1

(a)

For one combination of substitutions made, define the following variables:

Let $\vec{x}$ be an n-dimensional vector of 1s and 0s where x_i is 1 if the i^{th} substitution was made, and 0 if not.

Let $\vec{k}$ be an n-dimensional vector of weights, where k_i is the amount that making the i^{th} substitution contributes to the expression level.

Let c be the baseline expression level.

Let y be the expression level.

Then:

$$\vec{k} \cdot \vec{x} + c = y$$

If we have $n+1$ such equations, one for each combination, we can represent the linear system in the following matrix: (x_i will now be $x_{j,i}$ where j is the ordinal number of the combination/equation, and apply a similar transformation to the indexes of k and y)

$$\begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} & 1 \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \\ x_{n+1,1} & x_{n+1,2} & \cdots & x_{n+1,n} & 1 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n+1} \end{bmatrix}$$

We know the eqns are independent since the combinations of substitutions are distinct, and may not be scalar multiples of one another because all values are zero or one.

The problem is in the canonical $A\vec{x} = \vec{b}$ form. The A matrix is full-rank (number of rows equals number of columns), and you'll find exact solutions for $\vec{k}$ using Gaussian Elimination.

(b) Now we have n^2 equations rather than $n+1$ and the number of rows is greater than the number of columns.

Therefore the system is overdetermined. Given our overdetermined system $X\vec{k} = \vec{y}$, we find $\vec{k}$ which minimizes the sum of least-squares by solving the linear system $(X^T X)\vec{k} = (X^T \vec{y})$