Statistics 243: *class notes*

William J. De Meo

October 17, 1997

1 Memory Consideration

Suppose you want a QR decomposition of a 1000000 by 1500 matrix X. If you look at `top` and see that your program is using only 2.1% of the resources, you know you're moving data.

The X'X matrix is symmetric, so we should take advantage of this in our algorithm. For starters, to compute X'X itself, you should just do

$$(X^t X)_{ij} = dot(x_i, x_j)$$

i.e. the dot product of the ith column with the jth column.

The sum of squared residuals is a quadratic form. You can write an efficient algorithm to compute quadradic forms.

2 Cholesky Decomposition

If A is psd, there exists an upper triangular matrix U such that $U^t U = A$. Computing the decomposition:

$$\begin{aligned}
a_{ij} &= \sum_{k=1}^{i} u_{ki} u_{kj} \\
a_{ij} &= \sum_{k=1}^{i-1} u_{ki} u_{kj} + u_{ii} u_{ij} \\
u_{ij} &= \frac{1}{u_{ii}} (a_{ij} - \sum_{k=1}^{i-1} u_{ki} u_{kj} \\
u_{ii} &= (a_{ii} - \sum_{k=1}^{i-1} u_{ki}^2)^{1/2}
\end{aligned}$$

for i=1 to p

$a_{ii} = (a_{ii} - \sum_{k=1} i - 1 a_{ki}^2)^{1/2}$

for j=i+1 to p

$a_{ij} = (a_{ij} - \sum_{k=1}^{i-1} a_{ki} a_{kj}) / a_{ii}$

If any a_{ii} are close to zero, then we see that the algorithm fails and so A must not be positive semidefinite.

3 Regression

Do a Cholesky decomposition of $X^tX = U^tU$. Let $\theta = U\hat{\beta}$ Then $U^t\theta = X^ty$. Solve the l.t. system for θ, then solve the l.t. system $U^t\hat{\beta} = \theta$ for $\hat{\beta}$. When reading in your data, you would want to have the augmented matrix

$$[X : y]'[X : y]$$

Suppose X has mean θ variance Σ. Then consider U^t, which has

$$E(U^tX) = U^t\theta \text{ and } V(U^tX) = U^t\Sigma U$$

Suppose X has mean 0 and variance I. Then $X^* = U^tX$ has variance $\Sigma = U^tU$ So we can created correlated random variables with variance covariance structure Σ by starting with uncorrelated random variables in X, taking the Cholesky decomposition $\Sigma = U^tU$, and then apply U to X.