

1. Basic Math

- Working knowledge of integration such as integration by parts, area of simple geometry such as triangle and rectangle
- Solutions x_{\pm} of quadratic equation $ax^2 + bx + c = 0$: $x_{\pm} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$
- Working knowledge of matrix multiplication, which is essentially inner product of a row vector and a column vector
- Determinant of a 2×2 matrix $\mathbf{M} := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ $\det \mathbf{M} \equiv |\mathbf{M}| := ad - bc$
- Inverse of a 2×2 matrix \mathbf{M} : $\mathbf{M}^{-1} = \frac{1}{\det \mathbf{M}} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

2. Probability and Statistics

- Working knowledge of sample mean, unbiased sample variance, covariance, correlation, percentile by linear interpolation
- Bayes theorem $\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A) \mathbb{P}(A)}{\mathbb{P}(B)}$.
- Probabilities add up to one.
- Working knowledge to apply $\mathbb{E}(\cdot)$, $\mathbb{V}(\cdot)$, and $\mathbb{C}(\cdot)$ on both sides of the equation.

3. OLS

- Slope estimate of simple linear regression: $\hat{b} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$
- y intercept of simple linear regression: $\hat{a} = \bar{y} - \hat{b}\bar{x}$
- Normal distribution

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} \stackrel{d}{\sim} N \left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \sigma_e^2 \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right) & -\sigma_e^2 \left(\frac{\bar{X}}{\sum_{i=1}^n (X_i - \bar{X})^2} \right) \\ -\sigma_e^2 \left(\frac{\bar{X}}{\sum_{i=1}^n (X_i - \bar{X})^2} \right) & \sigma_e^2 \left(\frac{1}{\sum_{i=1}^n (X_i - \bar{X})^2} \right) \end{pmatrix} \right)$$

- Working knowledge of unbiased variance of the residuals and t statistic
 - When the sample size is more than 30, t statistic is close to the standard normal distribution.
 - The critical value at 2-sided 95% confidence level is 1.96.
 - For simple OLS, square of the t statistic is the F statistic.

5. Multiple linear regression estimates in general: $\hat{\beta} = (X'X)^{-1}X'y$.

6. Let $X = \begin{pmatrix} 1 & x \end{pmatrix}$ and $\beta = \begin{pmatrix} \alpha & \beta \end{pmatrix}'$.

$$X'X = \begin{pmatrix} 1' \\ x' \end{pmatrix} \begin{pmatrix} 1 & x \end{pmatrix} = \begin{pmatrix} 1'1 & 1'x \\ x'1 & x'x \end{pmatrix} = \begin{pmatrix} n & \sum x_t \\ \sum x_t & \sum x_t^2 \end{pmatrix}$$

$$(X'X)^{-1} = \frac{1}{n \sum x_t^2 - (\sum x_t)^2} \begin{pmatrix} \sum x_t^2 & -\sum x_t \\ -\sum x_t & n \end{pmatrix}$$

$$X'y = \begin{pmatrix} 1'y \\ x'y \end{pmatrix} = \begin{pmatrix} \sum y_t \\ \sum x_t y_t \end{pmatrix}$$

7. Therefore $(X'X)^{-1}X'y$ gives $\hat{\beta} = \begin{bmatrix} \hat{\alpha} & \hat{\beta} \end{bmatrix}'$:

$$\hat{\alpha} = \frac{\sum x_t^2 \sum y_t - \sum x_t \sum x_t y_t}{n \sum x_t^2 - (\sum x_t)^2} \quad \hat{\beta} = \frac{n \sum x_t y_t - \sum x_t \sum y_t}{n \sum x_t^2 - (\sum x_t)^2}$$

8. The affine form is $y = a + bx + e$. The mean square error is the variance of the error e , i.e.,

$$\sigma_e^2 := \mathbb{V}[e] = \text{MSE}.$$

9. OLS Algorithm for Model M

- Estimate the model by OLS: $\hat{\beta} = (X'X)^{-1}X'y$
- Compute the fitted values of y : $\hat{y} = X\hat{\beta}$
- Compute the residuals or "surprise": $\hat{u} = y - \hat{y}$
- Compute the residual sum of squares (RSS)

$$\text{SSE} \equiv \text{RSS} = \hat{u}'\hat{u} = \sum_{i=1}^n \hat{u}_i^2$$

e) The variance of the residuals is

$$\hat{\sigma}_u^2 = \frac{1}{n-K} \hat{u}'\hat{u}$$

f) Let $\Omega := (X'X)^{-1}$. The variance of $\hat{\beta}_i$ is

$$\mathbb{V}(\hat{\beta}_i) = \hat{\sigma}_u^2 \Omega_{ii}.$$

10. Statistical Inference

- For all $j = 1, 2, \dots, K$, the t test statistic for $\hat{\beta}_j$ is, given the null-hypothesized value β_j :

$$\frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}_u \sqrt{\Omega_{jj}}} \sim t_{n-K}$$

Here, $\Omega := (X'X)^{-1}$, and Ω_{jj} is the j -th diagonal element.

- The $\alpha\%$ significance level for β_j is, assuming two-tail test,

$$\hat{\beta}_j - q\hat{\sigma}_u \sqrt{\Omega_{jj}} \leq \beta_j \leq \hat{\beta}_j + q\hat{\sigma}_u \sqrt{\Omega_{jj}},$$

where q is the $(1 - \alpha/2)$ -th quantile of the t_{n-K} distribution.

4. Asset Pricing

1. Sharpe's ratio, M^2 , information ratio, Treynor's measure, and Jensen's alpha
2. SML, CML

5. Introductory Time Series and GARCH

1. Stationarity determination by characteristic equation
2. Working knowledge of VaR
3. $\omega = \gamma V_L$, where V_L is the long run variance
4. ARCH model for a process on the variance of u_t

$$\mathbb{V}(u_t) = \alpha_0 + \alpha_1 u_{t-1}^2$$

More generally, with $e_t \stackrel{d}{\sim} N(0, 1)$,

$$u_t = e_t \sqrt{\alpha_0 + \alpha_1 u_{t-1}^2}$$

5. GARCH(1,1)
6. Long-term variance in GARCH(1,1)

$$V_L = \frac{\omega}{\gamma} = \frac{\omega}{1 - \alpha - \beta}.$$