

Basic Mathematics for Portfolio Management

“If you stop at general math, you’re only going to make general math money.”

- Snoop Dogg



Statistics

- Variables x, y, z
- Constants a, b
- Observations $\{x_n, y_n, z_n \mid n=1, \dots, N\}$
- Mean

$\bar{x} = E\{x\}$ = mean or expected value of x

$$E\{a \cdot x + b \cdot y\} = a \cdot \bar{x} + b \cdot \bar{y}$$

$$\bar{x} = \frac{1}{N} \cdot \sum_{n=1}^N x_n$$

Variance

$$\sigma_x^2 = Var\{x\} = E\left\{(x - \bar{x})^2\right\}$$

$$Var\{a \cdot x\} = a^2 \cdot \sigma_x^2$$

$$\sigma_x^2 = \frac{1}{N-1} \cdot \sum_{n=1}^N (x_n - \bar{x})^2$$

Covariance

$$\sigma_{xy} = \text{Cov}\{x, y\} = E\{(x - \bar{x}) \cdot (y - \bar{y})\}$$

$$\sigma_{xy} = \sigma_{yx}$$

$$\text{Cov}\{x, x\} = \sigma_x^2$$

$$\text{Cov}\{a \cdot x + b \cdot y, z\} = a \cdot \sigma_{xz} + b \cdot \sigma_{yz}$$

$$\text{Var}\{a \cdot x + b \cdot y\} = \text{Cov}\{a \cdot x + b \cdot y, a \cdot x + b \cdot y\}$$

$$= a^2 \cdot \sigma_x^2 + b^2 \cdot \sigma_y^2 + 2ab \cdot \sigma_{xy}$$

$$\sigma_{xy} = \frac{1}{N-1} \cdot \sum_{n=1}^N (x_n - \bar{x}) \cdot (y_n - \bar{y})$$

Correlation

- Closely related to covariance:

$$\rho_{xy} \equiv \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$$

$$\sigma_x^2 = \sigma_x \cdot \sigma_x$$

$\sigma_x = \text{Standard Deviation}$

- Note that correlation is dimensionless, and that:

$$-1 \leq \rho_{xy} \leq 1$$

- This facilitates our intuition for the magnitude of a correlation.

A bit more on Correlation

- What is $\text{Corr}\{a \cdot x, y\}$?

$$\text{Corr}\{a \cdot x, y\} = \frac{\text{Cov}\{a \cdot x, y\}}{\sqrt{\text{Var}\{a \cdot x\}, \text{Var}\{y\}}} = \frac{a \cdot \sigma_{xy}}{a \cdot \sigma_x \cdot \sigma_y}$$

$$= \rho_{xy}$$

Regression

$$y_n = a + b \cdot x_n + \varepsilon_n = \hat{y}_n + \varepsilon_n$$

- This is an example of a linear model.
- The model is unbiased if:

$$\sum_{n=1}^N \varepsilon_n = 0$$

- How do we choose the best $\{a, b\}$?

Regression

- The best estimates for $\{a, b\}$ should minimize the sum of squared errors:

$$ESS = \sum_{n=1}^N \varepsilon_n^2 = \sum_{n=1}^N (y_n - \hat{y}_n)^2$$

- This corresponds to minimizing:

$$\sum_{n=1}^N \hat{y}_n^2 - 2 \cdot y_n \cdot \hat{y}_n$$

- Which leads to two equations:

$$\sum_{n=1}^N [\hat{y}_n - y_n] \cdot \left(\frac{\partial \hat{y}_n}{\partial a} \right) = 0$$

$$\sum_{n=1}^N [\hat{y}_n - y_n] \cdot \left(\frac{\partial \hat{y}_n}{\partial b} \right) = 0$$

Regression

- The first equation leads to an unbiased model.

$$\left(\frac{\partial \hat{y}_n}{\partial a} \right) = 1 \Rightarrow \sum_{n=1}^N [\hat{y}_n - y_n] = 0$$

- We can use this to solve for a :

$$y_n = a + b \cdot x_n + \varepsilon_n$$

$$\bar{y} = a + b \cdot \bar{x}$$

$$a = \bar{y} - b \cdot \bar{x}$$

Regression $y_n = a + bX_n + \varepsilon_n$

- The second equation relates the coefficient b to the sample covariance of x and y :

$$\left(\frac{\partial \hat{y}_n}{\partial b} \right) = x_n \Rightarrow \sum_{n=1}^N [\hat{y}_n - y_n] \cdot x_n = 0$$
$$\Rightarrow b = \frac{\text{Cov}\{x, y\}}{\text{Var}\{x\}}$$

- Putting this together:

$$y_n = \bar{y} + \left(\frac{\text{Cov}\{x, y\}}{\text{Var}\{x\}} \right) \cdot [x_n - \bar{x}] + \varepsilon_n$$

- The model \hat{y} is the best linear unbiased estimate (BLUE), given only the sample data.
- Before some examples, I need to introduce portfolio management notation.

Portfolio Management Definitions and Notation

$$\overbrace{t \quad \quad \quad t + \Delta t}$$

- Total rate of return:

$$trr \equiv \frac{\overset{\text{new price}}{\downarrow} p(t + \Delta t) + \overset{\text{dividend}}{\downarrow} d(t + \Delta t) - \overset{\text{old price}}{\downarrow} p(t)}{p(t)}$$

price and dividend notation are standard but we almost never focus on total rate of return.

- Risk-free return: i_F (Usually T-Bill return over period Δt)

- Excess return:

excess returns are a standard focus of investment theory

$$r \equiv trr - i_F$$

$$r_n = \text{stock} - n \text{ excess return}$$

Portfolios and Holdings

- A portfolio is a set of holdings.
- Holdings are value-weighted fractions:
 - Example: portfolio consists of 100 shares each of A and B. The price of A is \$50 and the price of B is \$25.

$$h_A = \frac{100 \cdot 50}{100 \cdot 50 + 100 \cdot 25} = 0.67$$

$$h_B = \frac{100 \cdot 25}{100 \cdot 50 + 100 \cdot 25} = 0.33$$

- Fully invested portfolios have holdings that sum to 1.
- Long-only portfolios have all holdings greater than or equal to 0.
- A benchmark is a particular portfolio, one outside the control of the portfolio manager. Examples include the S&P 500, the Russell 2000, and the MSCI EAFE.

Returns and Portfolios

- Portfolio returns: $r_P = \mathbf{h}_P^T \cdot \mathbf{r} = \sum_{n=1}^N h_p(n) \cdot r_n$
- Residual returns (defined via regression)

$$\underline{r_P} = \beta_P \cdot \underline{r_B} + \theta_P = \alpha_P + \beta_P \cdot r_B + \varepsilon_P$$

- Active returns (defined via subtraction)

$$r_P = r_B + \delta_P$$

- Active portfolio holdings (defined via subtraction)

$$\mathbf{h}_{PA} \equiv \mathbf{h}_P - \mathbf{h}_B$$

↪ overweights / underweights

Returns and Portfolios

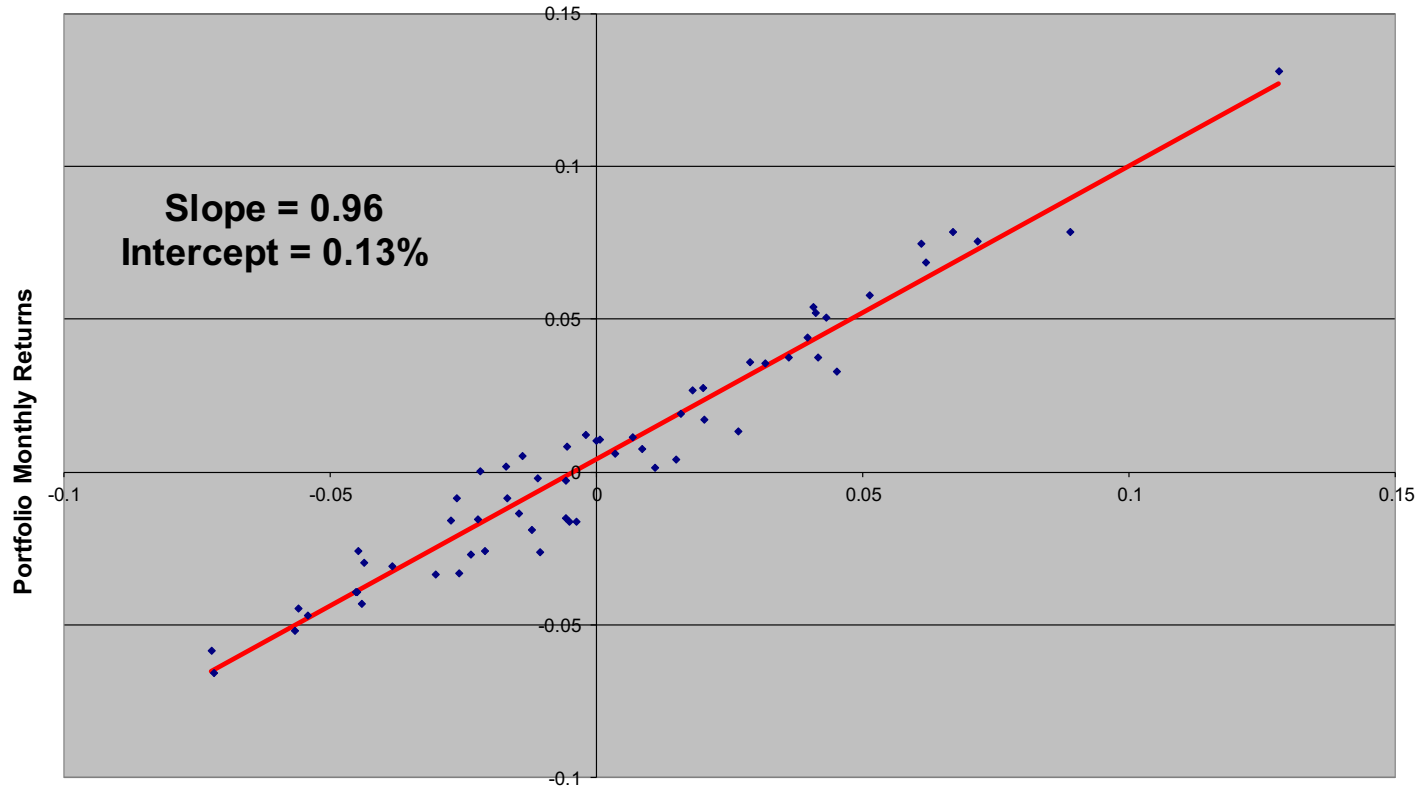
Returns	Variable	Mean	Variance
Excess (above risk-free)	r	f	σ^2
Residual	θ	α	ω^2
Active	δ	$\alpha + (\beta - 1) \cdot f_B$	$\omega^2 + (\beta - 1)^2 \cdot \sigma_B^2$

Portfolios:	\mathbf{h}_P	=	portfolio holdings
	\mathbf{h}_B	=	benchmark holdings
	\mathbf{h}_{PA}	=	active holdings
Covariance Matrix:	\mathbf{V}	=	covariance matrix ($\text{Cov}\{\mathbf{r}, \mathbf{r}\}$)

- Now to some portfolio management examples of regression.

Investment Example: Time-series Regression

Linear Regression



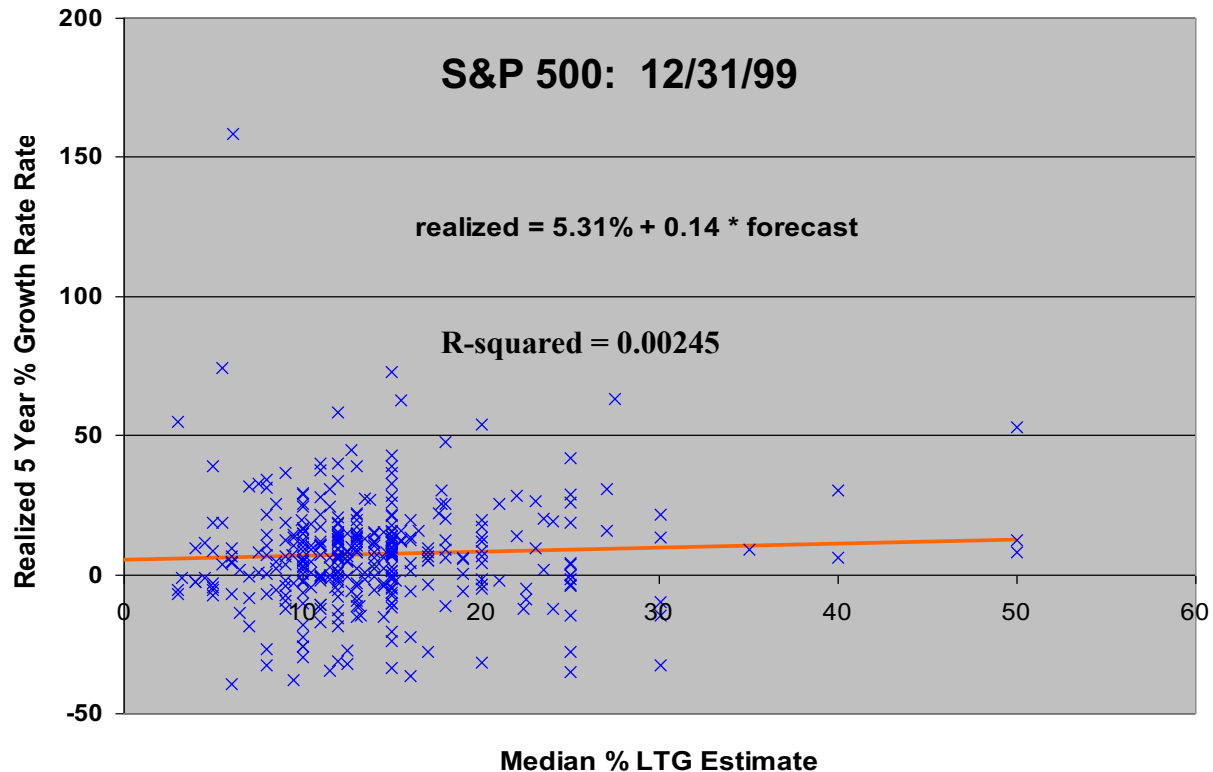
$$r_P(t) = \alpha_P + \beta_P \cdot r_B(t) + \varepsilon_P(t)$$

$$E\{r_P(t)\} = \alpha_P + \beta_P \cdot E\{r_B(t)\}$$

$$E\{r_P(t) | r_B(t)\} = \alpha_P + \beta_P \cdot r_B(t)$$

Benchmark Monthly Returns

Investment Example: Cross-sectional Regression



BLUE of y conditional on x

- More generally, the BLUE has the form:

$$E\{y|x\} = E\{y\} + Cov\{x, y\} \cdot Var^{-1}\{x\} \cdot (x - E\{x\})$$

- You can see how this directly relates to our prior regression result.
- We will apply this result more generally than just in the regression context. For example, our estimates of variances and covariances may improve upon sample estimates.

Matrices: Transpose and Inverse

- Transpose
 - Switch rows and columns

$$A_{mn}^T \equiv A_{nm}$$

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}^T = \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}$$

- Inverse (Assume A, B are square ($N \times N$) matrices)

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

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

$$\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I} = N \times N \text{ identity matrix}$$

$$(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}$$



Example: 2x2

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\mathbf{A}^{-1} = \frac{1}{(a_{11} \cdot a_{22} - a_{12} \cdot a_{21})} \cdot \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

- Verify that $\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Using Linear Algebra

- Portfolio Variance

$$\mathbf{h} = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_N \end{pmatrix}$$

$$\mathbf{V} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2N} \\ \vdots & & \ddots & \vdots \\ \sigma_{N1} & & & \sigma_N^2 \end{pmatrix}$$

N^2 elements

$$\sigma_P^2 = \text{Var}\{h_1 \cdot r_1 + h_2 \cdot r_2 + \dots + h_N \cdot r_N\}$$

$\frac{N(N-1)}{2}$ indep. elements

$$= h_1^2 \cdot \sigma_1^2 + \dots + h_N^2 \cdot \sigma_N^2 + 2h_1h_2 \cdot \sigma_{12} + 2h_1h_3 \cdot \sigma_{13} + \dots$$

N diagonal elements

$$= \mathbf{h}^T \cdot \mathbf{V} \cdot \mathbf{h}$$

Multivariate Regression

- The multivariate linear model is:

$$\mathbf{y} = \mathbf{X} \cdot \mathbf{b} + \boldsymbol{\varepsilon}$$

$$\boldsymbol{\Omega} = \text{Cov}\{\boldsymbol{\varepsilon}\}$$

- We can include an intercept as a column of \mathbf{X} .
- To generalize, we minimize the weighted sum of squared errors:

$$\boldsymbol{\varepsilon}^T \cdot \boldsymbol{\Omega}^{-1} \cdot \boldsymbol{\varepsilon}$$

- Our resulting estimates are:

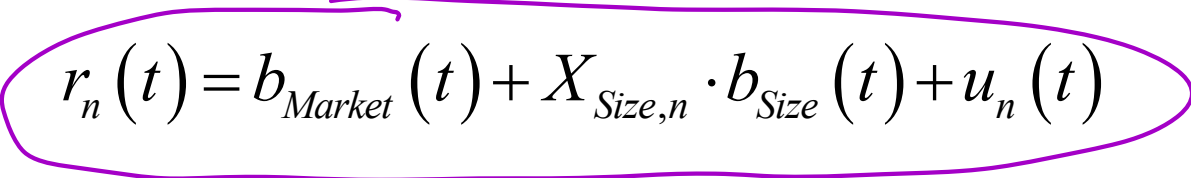
$$\mathbf{b} = \left[\mathbf{X}^T \cdot \boldsymbol{\Omega}^{-1} \cdot \mathbf{X} \right]^{-1} \cdot \mathbf{X}^T \cdot \boldsymbol{\Omega}^{-1} \cdot \mathbf{y}$$

Example: Simple 2-factor model

- Previously we saw a “beta regression” model:

$$r_P(t) = \alpha_P + \beta_P \cdot r_B(t) + \varepsilon_P(t)$$

- Now consider a different type of model:


$$r_n(t) = b_{Market}(t) + X_{Size,n} \cdot b_{Size}(t) + u_n(t)$$

where we have only two factors—a market factor and a measure of size—to model cross-sectional asset returns

What is X? What is b?

Market exposure

Size exposure

large stocks

sorted by market cap

Small stocks

$$\mathbf{X} = \begin{bmatrix} 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \\ 1 & -1 \\ \vdots & \vdots \\ 1 & -1 \end{bmatrix}$$
$$\mathbf{b} = [\mathbf{X}^T \cdot \mathbf{X}]^{-1} \cdot \mathbf{X}^T \cdot \mathbf{r}$$

(ignore GLS weighting)

All N stocks are in the market (all Market exposures = 1). $N/2$ stocks are large (size exposure = 1) and $N/2$ are small (size exposure = -1).

This is an odd definition of “the market.”

Basic Utility Function

- Mean/Variance

$$U = f_P - \lambda \cdot \sigma_P^2$$

$$= \mathbf{h}^T \cdot \mathbf{f} - \lambda \cdot \mathbf{h}^T \cdot \mathbf{V} \cdot \mathbf{h}$$

Portfolio Optimization

- Choose portfolio \mathbf{h} to maximize U .
- What does that mean? We must take the derivative of U with respect to each of the N elements of \mathbf{h} . That leads to N equations in N unknowns.

$$\frac{\partial U}{\partial h_n} = 0$$

$$\frac{\partial U}{\partial \mathbf{h}^T} = 0$$

More Hints

- Hint 2: Try out 2x2 or 3x3 examples.
- Hint 3: To start, try working with individual elements. For example, I will use shorthand like:

$$\frac{\partial}{\partial \mathbf{h}^T}$$

- This says to take derivatives with respect to each element of \mathbf{h} . This should lead to N separate equations.

- Example:
$$\frac{\partial}{\partial \mathbf{h}^T} (\mathbf{h}^T \cdot \mathbf{V} \cdot \mathbf{h}) = 2\mathbf{V} \cdot \mathbf{h}$$

Portfolio Construction with Constraints

- We usually need to include constraints in our portfolio optimization.

- For example, the full investment constraint:

$$\sum_{n=1}^N h_n = 1$$

$$e = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$
$$h^T \cdot e = \sum h_n$$

- There is nothing in the mathematical optimization problem to guarantee the holdings sum to 1 unless we explicitly add it.
- We can add linear constraints using Lagrange multiplier:

$$U: \left[\mathbf{h}^T \cdot \mathbf{f} - \lambda \cdot \mathbf{h}^T \cdot \mathbf{V} \cdot \mathbf{h} \right] \Rightarrow \left[\mathbf{h}^T \cdot \mathbf{f} - \lambda \cdot \mathbf{h}^T \cdot \mathbf{V} \cdot \mathbf{h} + c \cdot (\mathbf{h}^T \cdot \mathbf{e} - 1) \right]$$

- We have added a new variable, c . Now when we optimize U , we take derivatives with respect to \mathbf{h} and c . Setting derivative with respect to c to zero leads to the constraint.