FINM 36700 — Lecture 2 TA Review: Linear Factor Decomposition

---

1) Linear Factor Decomposition (LFD)

The core setup is:

$$\boxed{\tilde{r}_{i,t} = \alpha_i + \boldsymbol{\beta}_i^\top \tilde{\mathbf{x}}_t + \epsilon_{i,t}}$$

where:

- $\tilde{r}_{i,t}$ = excess return of asset/portfolio i
- $\tilde{\mathbf{x}}_t$ = vector of factor returns
- $\boldsymbol{\beta}_i$ = factor exposures (betas)
- α_i = mean return not explained by the chosen factors
- $\epsilon_{i,t}$ = residual / idiosyncratic component

The purpose is to **decompose an asset's return into interpretable building blocks**:

$$\boxed{\text{Return} = \text{Factor-driven component} + \text{Residual component}}$$

LFD is fundamentally a **statistical projection**, not an asset-pricing or no-arbitrage model (like CAPM or Fama-French). A large $\hat{\alpha}$ does not automatically imply managerial skill; it may simply mean that important factors were omitted (omitted variable bias).

What can the factors be?

Factors may represent broad markets/asset classes, sectors/industries, regions, styles (value, momentum, size, quality), or macro/FX exposures. The factors should be chosen based on the specific economic question being asked.

---

2) OLS Estimation

Write the regression in matrix form:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon}$$

where $\boldsymbol{\theta} = [\alpha, \boldsymbol{\beta}^\top]^\top$. OLS chooses $\hat{\boldsymbol{\theta}}$ to minimize $\sum_t \hat{\epsilon}_t^2$. The solution is:

$$\boxed{\hat{\boldsymbol{\theta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}}$$

The key requirement for the projection is that $\mathbf{X}$ has full column rank:

$$\boxed{\operatorname{rank}(\mathbf{X}) = k + 1}$$

so that $\mathbf{X}^\top \mathbf{X}$ is invertible. Perfect multicollinearity prevents the factor exposures from being uniquely identified.

Interpretation

- $\boxed{\hat{\alpha}}$: Estimated average non-factor return.
- $\boxed{\hat{\beta}_j}$: Marginal sensitivity to factor j, holding other factors fixed.
- $\boxed{\hat{\epsilon}_t}$: What the factor model did not explain.
- $\boxed{R^2 = 1 - \dfrac{\operatorname{Var}(\hat{\boldsymbol{\epsilon}})}{\operatorname{Var}(r_i)}}$: Share of variation in the asset return explained by the chosen factors.

> Highly correlated regressors make $\mathbf{X}^\top \mathbf{X}$ ill-conditioned, leading to unstable beta estimates (inflated standard errors) even if the regression is technically estimable. Remedies include factor grouping, orthogonalization, or penalized regression (Ridge/Lasso).

Geometric interpretation of OLS:

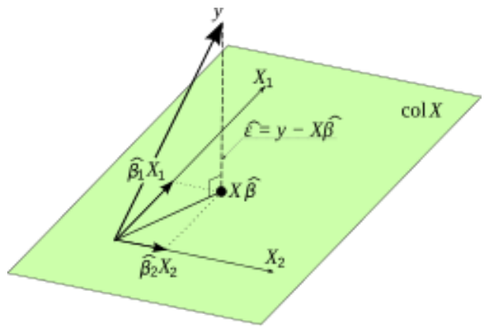


Figure: OLS projects the target vector onto the column space of the regressors.

---

3) What LFD is Useful For

The same regression can answer different questions depending on how we interpret the RHS factors. The main use cases are:

$$\boxed{\text{1. Performance Attribution}} \quad \boxed{\text{2. Hedging}} \quad \boxed{\text{3. Tracking / Replication}} \quad \boxed{\text{4. Risk-Model Dimension Reduction}}$$

The regression itself is identical. What changes is **what we do with the betas and residuals**.

---

4) Performance Attribution

Suppose $\tilde{r}_i = \alpha + \boldsymbol{\beta}^\top \tilde{\mathbf{x}} + \epsilon$. The expected return decomposes into:

$$\boxed{\mathbb{E}[\tilde{r}_i] = \alpha + \boldsymbol{\beta}^\top \mathbb{E}[\tilde{\mathbf{x}}]}$$

This lets us ask: *Was the portfolio's return simply compensation for factor exposures, or was there return left unexplained by those factors?*

If the chosen factors are broad investable benchmarks, α is the mean return that cannot be reproduced from those factor exposures. However, $\boxed{\alpha \text{ is always model-dependent.}}$ Changing the factor set will change α.

Case Studies: Performance Attribution for Berkshire Hathaway, PIMCO, and the Quantum Fund

These screenshots compare historical performance with regression-based factor attribution. The figures report returns in excess of cash; the sample periods differ across managers. Each factor's contribution to average return is its estimated beta multiplied by its average factor return. Alpha is the average return left unexplained by the selected factors, while R^2 measures the share of return variation explained by the regression.

Berkshire Hathaway — January 1977–May 2016

1/1977–5/2016[8]	Average Return	Volatility	Sharpe Ratio	Annual Outperformance	Information Ratio
Berkshire Hathaway	17.6%	23.6%	0.74	10.6%	0.49
US Equities[9]	6.9%	15.5%	0.45	–	–

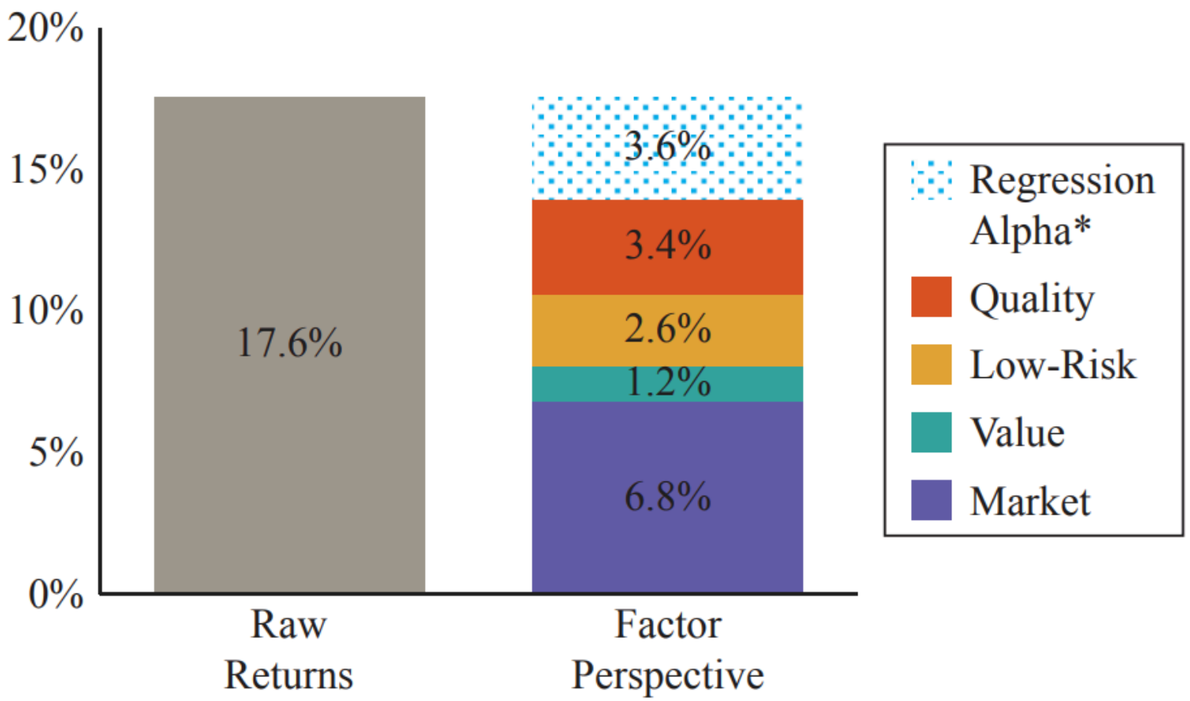
Sources: AQR, CRSP (for BRK data), Kenneth French Data Library (for Equities, which are CRSP cap-weighted returns; and risk-free rate, which is one-month Treasuries). For consistency, we've chosen the CRSP cap-weighted index to represent US Equities throughout this article. Returns are excess of cash throughout this article.

Performance: Berkshire Hathaway's average annual excess return was **17.6%**, compared with **6.9%** for US equities. Its volatility was **23.6%** and its Sharpe ratio was **0.74**, versus **15.5%** and **0.45** for US equities.

Panel A: Buffett's Factors

	Alpha (ann'l)	Market	Value	Low-Risk	Quality	R^2
Coefficient	3.6%	0.98	0.43	0.23	0.63	28%
T-Stat	1.03	12.41	3.60	2.64	4.64	

Panel B: Return Attributions Based on Regression Results, January 1977–May 2016



Factor attribution: The chart attributes approximately **6.8%** to the market, **1.2%** to value, **2.6%** to low-risk, and **3.4%** to quality, with **3.6% annual alpha**. The regression explains **28%** of return variation.

PIMCO Total Return Fund — January 1994–September 2014

1/1994–9/2014[17]	Average Return	Volatility	Sharpe Ratio	Annual Outperformance	Information Ratio
PIMCO Total Return Fund	4.3%	4.1%	1.05	1.4%	0.83
Barclays US Aggregate	2.9%	3.6%	0.80	–	–

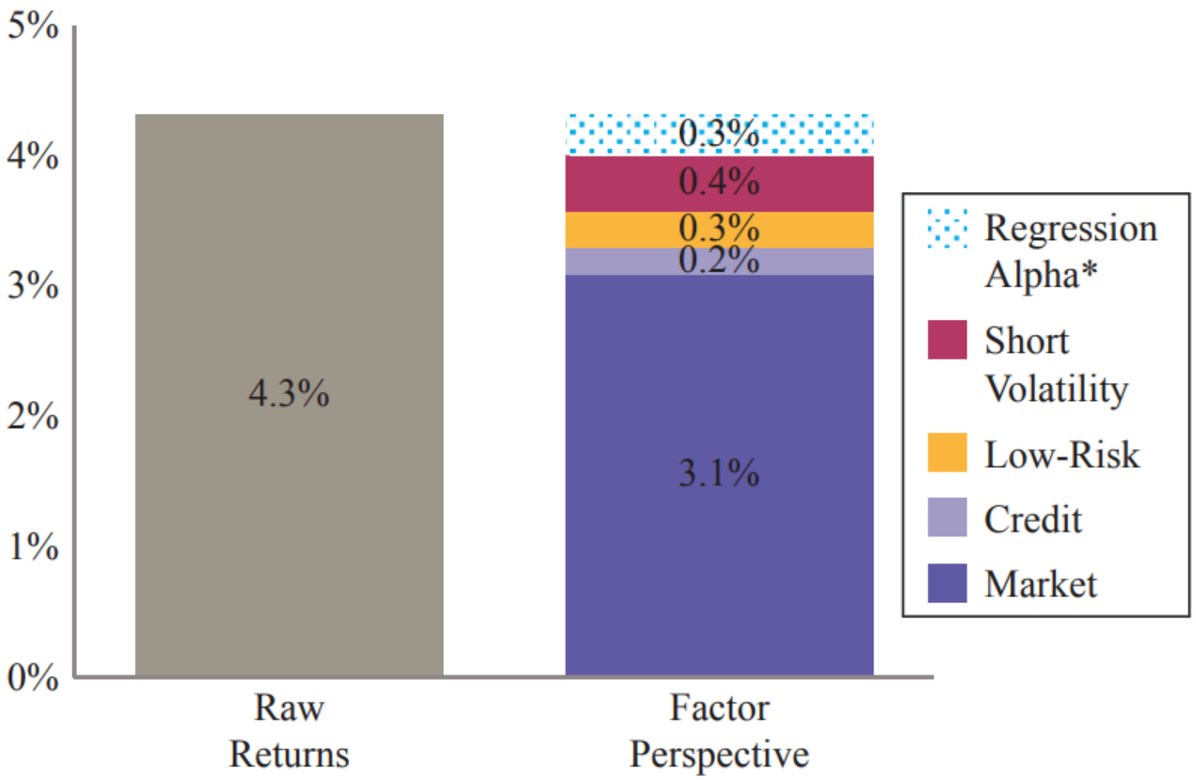
Sources: AQR, CRSP. Risk-free rate is one-month Treasury bills.

Performance: PIMCO's average annual excess return was **4.3%**, compared with **2.9%** for the Barclays US Aggregate benchmark. Its Sharpe ratio was **1.05**, versus **0.80** for the benchmark, with annual outperformance of **1.4%**.

Panel A: Gross Factors

	Alpha (ann'l)	Market	Credit	Low-Risk	Short Volatility	R^2
Coefficient	0.3%	1.06	0.06	0.07	0.04	88%
T-Stat	0.94	41.01	6.02	7.98	4.53	

Panel B: Return Attributions Based on Regression Results, January 1994–September 2014



Factor attribution: The chart attributes approximately **3.1%** to the bond-market factor, **0.2%** to credit, **0.3%** to low-risk, and **0.4%** to short volatility, leaving **0.3% annual alpha**. The regression explains **88%** of return variation.

George Soros' Quantum Fund — March 1985–May 2004

THE QUANTUM FUND—EQUITIES, TREND (EVERYWHERE), AND FUNDAMENTAL CURRENCY TRADING

"We try to catch new trends early and in later stages we try to catch trend reversals. Therefore, we tend to stabilize rather than destabilize the market. We are not doing this as a public service. It is our style of making money."

—George Soros[26]

3/1985–5/2004[27]	Average Return	Volatility	Sharpe Ratio
Quantum Fund	20.2%	23.1%	0.88
US Stock Market	7.8%	15.8%	0.49

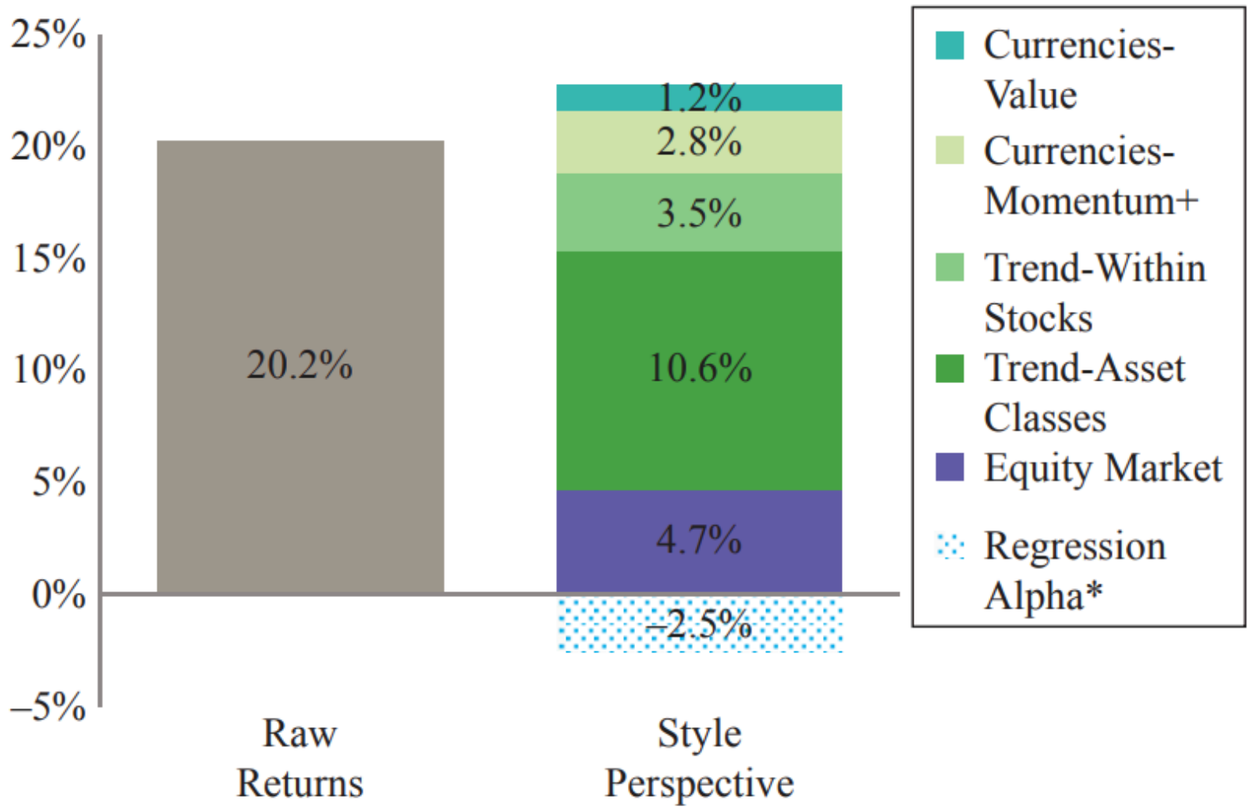
Sources: AQR, HFR. Risk-free rate is 1-month Treasuries. For consistency, we've chosen the CRSP cap-weighted index to represent US Equities throughout this article.

Performance: The Quantum Fund's average annual excess return was **20.2%**, compared with **7.8%** for the US stock market. Its volatility was **23.1%** and its Sharpe ratio was **0.88**, versus **15.8%** and **0.49** for the market.

Panel A: Soros Factors

	Alpha (Ann'l)	Market	Trend		Currencies		R^2
			Within Stocks	Asset Classes	Momentum	Value	
Coefficient	–2.5%	0.60	0.32	0.59	0.56	0.29	38%
T-Stat	–0.53	7.50	3.69	5.29	4.49	2.31	

Panel B: Return Attributions Based on Regression Results, March 1985–May 2004



Factor attribution: The chart attributes approximately **4.7%** to the equity market, **10.6%** to trend across asset classes, **3.5%** to trend within stocks, **2.8%** to currency momentum, and **1.2%** to currency value. These contributions exceed the fund's average return, resulting in **–2.5% annual alpha**. The regression explains **38%** of return variation.

---

5) Information Ratio (IR)

The factor-adjusted return is $\alpha + \epsilon_t$. Its expected value is α, and its volatility is σ_ϵ. Therefore:

$$\boxed{IR = \frac{\alpha}{\sigma_\epsilon}}$$

The Information Ratio is simply the **Sharpe ratio of the residual (factor-adjusted) return**. It is the primary metric for judging whether a manager generates meaningful return *after controlling for systematic exposures*.

The Information Ratio here measures the tradeoff between obtaining extra mean return at the cost of taking on tracking error from the target portfolio.

---

6) Hedging: The Basic Two-Asset Problem

Suppose we are long $1 of asset i and hedge with h dollars of asset j$. The hedged return (the basis) is:

$$\boxed{\epsilon_t = r_{i,t} - h r_{j,t}}$$

Its variance (basis risk) is:

$$\mathrm{Var}(\epsilon) = \sigma_i^2 + h^2 \sigma_j^2 - 2h\sigma_i\sigma_j\rho_{ij}$$

To minimize basis risk, differentiate with respect to h and set to zero:

$$\frac{\partial}{\partial h}\mathrm{Var}(\epsilon) = 2h\sigma_j^2 - 2\sigma_i\sigma_j\rho_{ij} = 0$$

Solving for h yields the optimal hedge ratio:

$$\boxed{h^* = \frac{\mathrm{Cov}(r_i, r_j)}{\mathrm{Var}(r_j)} = \rho_{ij}\frac{\sigma_i}{\sigma_j}}$$

This is exactly the **regression beta** of i on j.

- If $\rho_{ij} < 0$, then $h^* < 0$, meaning we go **long** the hedge asset to offset risk.

---

7) Hedging Multiple Factors & Market-Neutral Positions

For multiple hedges, $r_i = \boldsymbol{\beta}^\top \mathbf{x} + \epsilon$. The OLS beta vector minimizes $\mathrm{Var}(r_i - \boldsymbol{\beta}^\top \mathbf{x})$. Thus, β_j is the optimal hedge ratio against factor j, *conditional on the other factors*. Multivariate beta is vastly more useful than pairwise correlation because it isolates marginal exposure.

Market-Neutral Position: If $\tilde{r}_i = \alpha + \beta_m \tilde{r}_m + \epsilon$, removing market exposure leaves $\tilde{r}_i - \beta_m \tilde{r}_m = \alpha + \epsilon$. The hedged position has mean α and volatility σ_ϵ, so its performance is naturally evaluated by $IR = \alpha / \sigma_\epsilon$.

---

8) Tracking vs. Hedging

Feature	Hedging	Tracking / Replication
Objective	**Remove** factor exposure	**Reproduce** factor exposure
Position	We own i, want to eliminate risk from $\mathbf{x}$	We cannot own i, want to mimic it using $\mathbf{x}$
Trade	Sell $\boldsymbol{\beta}$ (Short the factors)	Buy $\boldsymbol{\beta}$ (Long the factors)
Primary Metric	Basis Risk (σ_ϵ)	Tracking Error (σ_ϵ)
Alpha Target	Large α (skill)	$\alpha \approx 0$ (perfect replication)

---

9) The Intercept: Include α or Not?

This is a critical practical choice depending on the objective:

- **Include α (Unbudgeted / Attribution / Hedging):** Allows $\boldsymbol{\beta}$ to focus purely on matching return **variation** (covariance). This prevents noisy sample mean estimates from distorting the betas.
- **Omit α (Budgeted Replication):** Enforces "span" (level matching). Forces $\boldsymbol{\beta}$ to match both variation and the mean. *Warning:* This can degrade the R^2 fit if sample means are noisy, because it forces the betas to "predict" mean differences that may not persist out-of-sample.

> In most cases, if you are hedging or attributing, include α.

---

10) LFD for Dimension Reduction in Risk Models

The Problem

A portfolio of $N = 442$ S&P 500 stocks has 442 return series. Directly estimating the full covariance matrix requires roughly $\frac{N(N-1)}{2} \approx 97,000$ pairwise covariances. A risk model cannot reliably estimate this many parameters from a standard time series.

Solution: Factor Structure

We replace the massive covariance matrix with a regression on a handful of $K \ll N$ tradable sector ETFs (e.g., $K = 11$ sectors):

$$\boxed{r_{i,t} = \alpha_i + \boldsymbol{\beta}_i^\top \mathbf{f}_t + \epsilon_{i,t}}$$

The covariance structure decomposes into:

$$\boxed{\Sigma_r = B\Sigma_f B^\top + \Sigma_\epsilon}$$

where B is the $N \times K$ matrix of betas, Σ_f is the $K \times K$ factor covariance, and Σ_ϵ is the $N \times N$ residual covariance.

Why Dimension Reduction Works

1. **Diagonal Residuals:** Because residual correlations are near zero, Σ_ϵ can be safely approximated as a **diagonal matrix**.
2. **Dimensionality:** The problem of estimating 97,000 covariances is reduced to estimating $11 \times 12/2 = 66$ factor covariances, $442 \times 11 = 4,862$ betas, and 442 residual variances.

Key Formulas:

1. **LFD:** $\tilde{r}_i = \alpha + \boldsymbol{\beta}^\top \tilde{\mathbf{x}} + \epsilon$
2. **Optimal Hedge:** $h^* = \beta = \rho_{ij}\frac{\sigma_i}{\sigma_j}$
3. **Information Ratio:** $IR = \frac{\alpha}{\sigma_\epsilon}$
4. **Factor Covariance:** $\Sigma_r \approx B\Sigma_f B^\top + D$ (where D is diagonal)