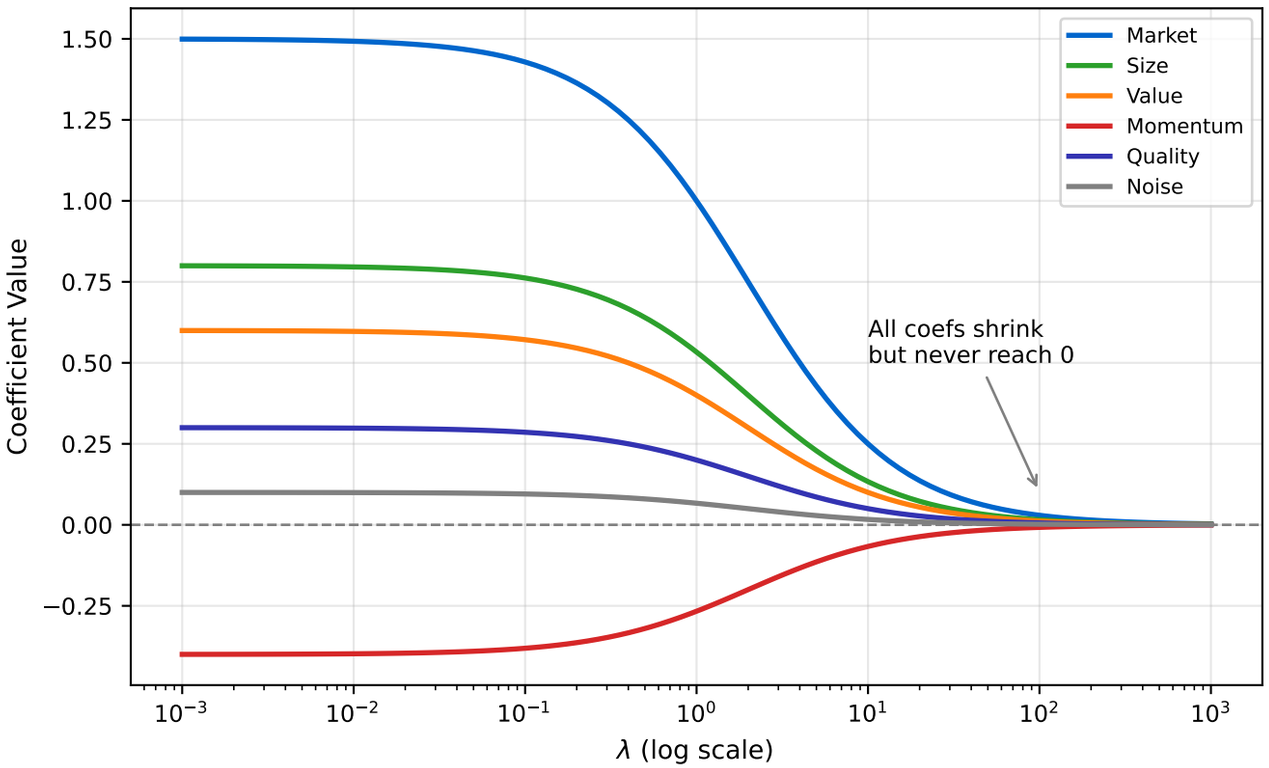
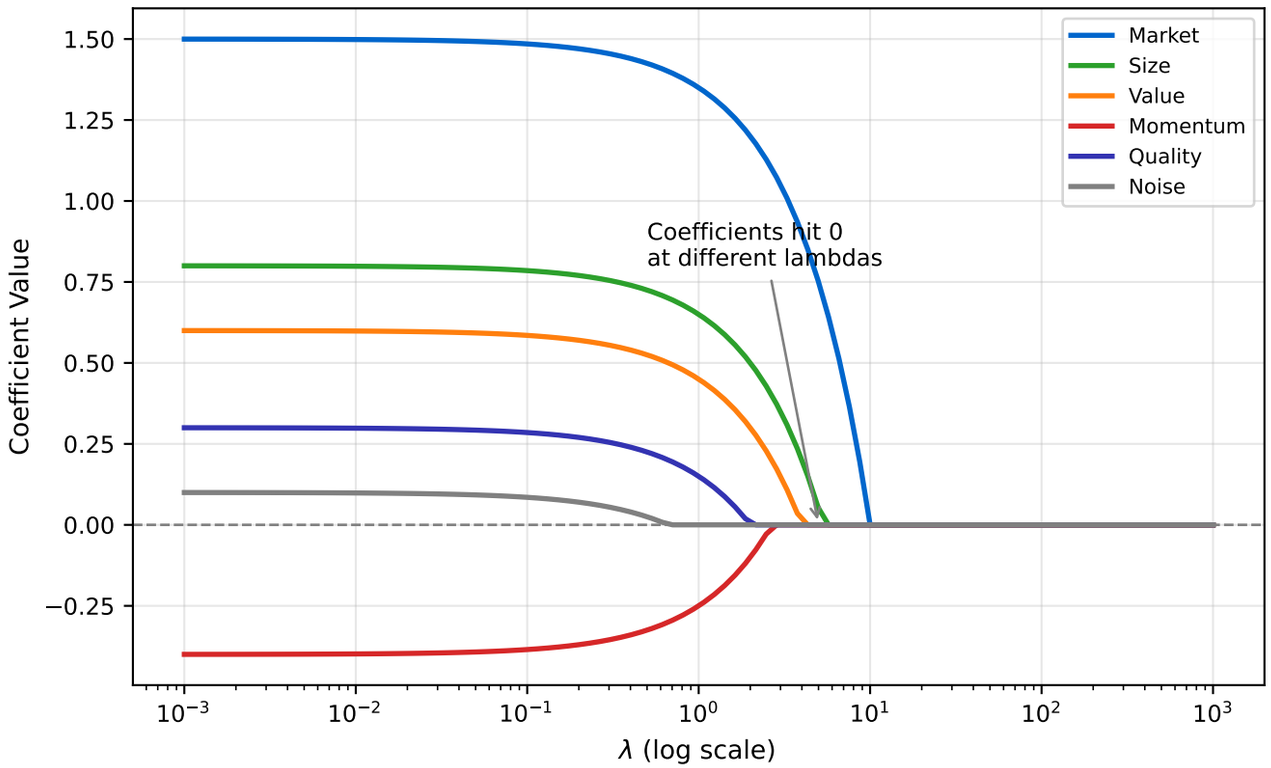
Regularization Paths: Coefficients vs Lambda

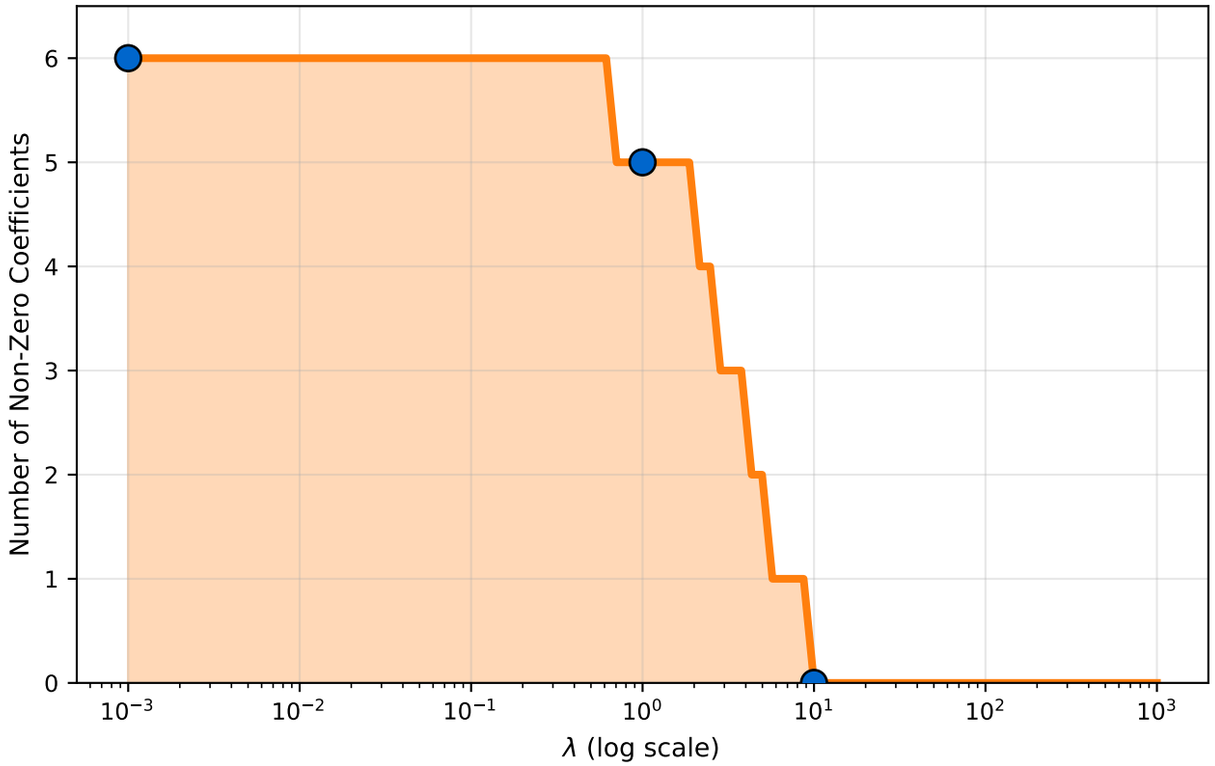
Ridge Path: Smooth Shrinkage



Lasso Path: Sparse Selection



Lasso: Number of Selected Features



Ridge vs Lasso Summary

```
RIDGE vs LASSO COEFFICIENT PATHS

RIDGE (L2):
- Coefficients shrink smoothly toward 0
- Never exactly reach 0
- All features stay in model
- Good when: all features are relevant

LASSO (L1):
- Coefficients hit 0 at different lambda values
- Automatic feature selection
- Sparse solutions
- Good when: some features are irrelevant

CHOOSING LAMBDA:
- Small lambda: More features, potential overfit
- Large lambda: Fewer features, potential underfit
- Use cross-validation to find optimal lambda

ELASTIC NET (combines both):
- Mix of L1 and L2 penalties
- sklearn: ElasticNet(alpha=1.0, l1_ratio=0.5)
- l1_ratio: 0 = Ridge, 1 = Lasso, 0.5 = balanced

In Finance:
- Ridge: Factor models (keep all factors)
- Lasso: Variable selection (find key factors)
```