Multivariate Break-Even Correlation Thresholds

Yannick Kälber • 24 Feb 2026

As we all know, backtesting is not a research tool, but the very end of your research pipeline. If you want to evaluate if a given set of signals is predictive for returns, you can do this more clearly and directly by regressing returns on the signals or measuring their correlations. But “*how strong*” do those correlations need to be for the signals to be “*good enough*”? And what about their interaction?

Linear Model

Say we have $n \in \mathbb{N}_{>0}$ signals collected in a vector $\mathbf{x} \in \mathbb{R}^n$ and a scalar return $r \in \mathbb{R}$ we try to predict. We model the return as a linear function of the signals:

$$r = \alpha + \boldsymbol{\beta}^\top \mathbf{x} + \epsilon \tag{1}$$

$$r = \alpha + \boldsymbol{\beta}^\top \mathbf{x} + \sigma_\varepsilon \varepsilon \tag{2}$$

$\alpha = \mu_r - \boldsymbol{\beta}^\top \boldsymbol{\mu}_x \in \mathbb{R}$ is the regression intercept, $\boldsymbol{\beta} \in \mathbb{R}^n$ is the vector of slope coefficients, and the residual $\epsilon \in \mathbb{R}$ is a mean-zero random variable with variance $\sigma_\varepsilon^2 \in \mathbb{R}_{>0}$. We can decompose ϵ in (1) into a scaled unit-variance residual term $\sigma_\varepsilon \varepsilon \in \mathbb{R}$, where ε has unit variance, yielding (2). We assume $\mathrm{Cov}(\mathbf{x}, \varepsilon) = \mathbf{0}$, which we use throughout. Let $\mu_r = \mathbb{E}[r] \in \mathbb{R}$ denote the unconditional expected return and $\boldsymbol{\mu}_x = \mathbb{E}[\mathbf{x}] \in \mathbb{R}^n$ the signal mean vector.

From the Model to Correlations

The marginal correlation between r and a signal x_i is $\rho_i \in [-1, 1]$, defined as $\rho_i = \mathrm{Cov}(r, x_i)/(\sigma_r \sigma_{x_i})$, where $\sigma_{x_i} \in \mathbb{R}_{>0}$ is the standard deviation of the i-th signal and $\sigma_r \in \mathbb{R}_{>0}$ is the standard deviation of the returns r. We collect the n marginal standard deviations σ_{x_i} in the diagonal matrix $\mathbf{D}_x = \mathrm{diag}(\sigma_{x_1}, \ldots, \sigma_{x_n}) \in \mathbb{R}^{n \times n}$ and the n correlations ρ_i in the vector $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_n)^\top \in \mathbb{R}^n$. Dividing each component of $\mathrm{Cov}(\mathbf{x}, r)$ by the corresponding term $\sigma_r \sigma_{x_i}$ amounts to pre-multiplying by $(\sigma_r \mathbf{D}_x)^{-1} = \frac{1}{\sigma_r} \mathbf{D}_x^{-1}$. Substituting (2) into $\mathrm{Cov}(\mathbf{x}, r)$ and carrying through:

$$\boldsymbol{\rho} = \frac{1}{\sigma_r}\mathbf{D}_x^{-1}\mathrm{Cov}(\mathbf{x}, r) \tag{3}$$

$$= \frac{1}{\sigma_r}\mathbf{D}_x^{-1}\mathrm{Cov}(\mathbf{x}, \alpha + \boldsymbol{\beta}^\top\mathbf{x} + \sigma_\varepsilon\varepsilon) \tag{4}$$

$$= \frac{1}{\sigma_r}\mathbf{D}_x^{-1}\left[\mathrm{Cov}(\mathbf{x}, \alpha) + \mathrm{Cov}(\mathbf{x}, \boldsymbol{\beta}^\top\mathbf{x}) + \sigma_\varepsilon\mathrm{Cov}(\mathbf{x}, \varepsilon)\right] \tag{5}$$

$$= \frac{1}{\sigma_r}\mathbf{D}_x^{-1}\left[\mathbf{0} + \boldsymbol{\Sigma}_x\boldsymbol{\beta} + \mathbf{0}\right] \tag{6}$$

$$= \frac{1}{\sigma_r}\mathbf{D}_x^{-1}\boldsymbol{\Sigma}_x\boldsymbol{\beta} \tag{7}$$

$$= \frac{1}{\sigma_r}\mathbf{R}_x\mathbf{D}_x\boldsymbol{\beta} \tag{8}$$

In (3) we write the vector form of the correlation definition, and in (4) we substitute (2) for r. In (5) we use the linearity of covariance. In (6) we use three facts: $\mathrm{Cov}(\mathbf{x}, \alpha) = \mathbf{0}$ because α is a constant; $\mathrm{Cov}(\mathbf{x}, \boldsymbol{\beta}^\top\mathbf{x}) = \boldsymbol{\Sigma}_x\boldsymbol{\beta}$ where $\boldsymbol{\Sigma}_x = \mathrm{Var}(\mathbf{x}) \in \mathbb{R}^{n\times n}$ is the covariance matrix of the signals; and $\mathrm{Cov}(\mathbf{x}, \varepsilon) = \mathbf{0}$ by our assumption. In (7) we collect terms, and in (8) we simplify $\mathbf{D}_x^{-1}\boldsymbol{\Sigma}_x$ by recalling how the covariance matrix decomposes into correlations and standard deviations: the signal correlation matrix $\mathbf{R}_x \in \mathbb{R}^{n\times n}$ is defined by $\boldsymbol{\Sigma}_x = \mathbf{D}_x\mathbf{R}_x\mathbf{D}_x$, so pre-multiplying by $\mathbf{D}_x^{-1}$ gives $\mathbf{D}_x^{-1}\boldsymbol{\Sigma}_x = \mathbf{R}_x\mathbf{D}_x$.

From Correlations to Betas

For stating a signal evaluation criterion in the next step, we need $\boldsymbol{\beta}$ to be expressed in terms of $\boldsymbol{\rho}$. We obtain this by inverting (8), multiplying both sides on the left by $\sigma_r\mathbf{D}_x^{-1}\mathbf{R}_x^{-1}$:

$$\sigma_r\mathbf{D}_x^{-1}\mathbf{R}_x^{-1}\boldsymbol{\rho} = \mathbf{D}_x^{-1}\mathbf{R}_x^{-1}\mathbf{R}_x\mathbf{D}_x\boldsymbol{\beta} \tag{9}$$

$$= \mathbf{D}_x^{-1}\mathbf{D}_x\boldsymbol{\beta} \tag{10}$$

$$= \boldsymbol{\beta} \tag{11}$$

In (9) we multiply both sides of (8) by $\sigma_r\mathbf{D}_x^{-1}\mathbf{R}_x^{-1}$, where the $\frac{1}{\sigma_r}$ on the right cancels with σ_r. In (10) the product $\mathbf{R}_x^{-1}\mathbf{R}_x = \mathbf{I}$ reduces to an identity, and in (11) $\mathbf{D}_x^{-1}\mathbf{D}_x = \mathbf{I}$ reduces to an identity as well. Reading from right to left:

$$\boldsymbol{\beta} = \sigma_r\mathbf{D}_x^{-1}\mathbf{R}_x^{-1}\boldsymbol{\rho} \tag{12}$$

This is the multivariate generalisation of the well-known univariate identity $\beta = \rho\sigma_r/\sigma_x$ linking the regression slope to the correlation coefficient, where the inverse correlation matrix $\mathbf{R}_x^{-1}$ adjusts for cross-correlations among the signals by isolating each signal's unique contribution conditional on the others.

Signal Evaluation Criterion

Finally, we state what it means for the correlations $\boldsymbol{\rho}$ of a set of signals to be "*good enough*". We require that, at a signal level $\mathbf{k} \in \mathbb{R}^n$ standard deviations from the mean, i.e. at $\mathbf{x} = \boldsymbol{\mu}_x + \mathbf{D}_x\mathbf{k}$, the corresponding absolute expected return exceeds a trading cost threshold $c \in \mathbb{R}_{>0}$:

$$\left|\mathbb{E}[r \mid \mathbf{x} = \boldsymbol{\mu}_x + \mathbf{D}_x\mathbf{k}]\right| > c \tag{13}$$

$$\left|\alpha + \boldsymbol{\beta}^\top(\boldsymbol{\mu}_x + \mathbf{D}_x\mathbf{k})\right| > c \tag{14}$$

$$\left|\alpha + \boldsymbol{\beta}^\top\boldsymbol{\mu}_x + \boldsymbol{\beta}^\top\mathbf{D}_x\mathbf{k}\right| > c \tag{15}$$

$$\left|\mu_r + \boldsymbol{\beta}^\top\mathbf{D}_x\mathbf{k}\right| > c \tag{16}$$

$$\left|\mu_r + (\sigma_r\mathbf{D}_x^{-1}\mathbf{R}_x^{-1}\boldsymbol{\rho})^\top\mathbf{D}_x\mathbf{k}\right| > c \tag{17}$$

$$\left|\mu_r + \sigma_r\boldsymbol{\rho}^\top\mathbf{R}_x^{-1}\mathbf{k}\right| > c \tag{18}$$

In (13) we state the criterion in general terms: the conditional expected return, evaluated at a joint signal realisation $\mathbf{k}$ standard deviations from the mean, must exceed the threshold c in absolute value. In (14) we substitute $\mathbb{E}[r \mid \mathbf{x}] = \alpha + \boldsymbol{\beta}^\top\mathbf{x}$ from (2). In (15) we distribute $\boldsymbol{\beta}^\top$ over the sum. In (16) we apply the OLS identity $\alpha + \boldsymbol{\beta}^\top\boldsymbol{\mu}_x = \mu_r$, absorbing the signal means into the unconditional expected return. In (17) we substitute (12) for $\boldsymbol{\beta}$. In (18) we transpose, using the symmetry of $\mathbf{R}_x^{-1}$ and $\mathbf{D}_x^{-1}$, and cancel $\mathbf{D}_x^{-1}\mathbf{D}_x = \mathbf{I}$. The absolute value reflects that the signals can be profitable in either direction (long or short).

Since the term $\boldsymbol{\rho}^\top\mathbf{R}_x^{-1}\mathbf{k}$ appears repeatedly throughout the rest of the article, we name it **combined signal strength** $\eta \in \mathbb{R}$:

$$\eta = \boldsymbol{\rho}^\top\mathbf{R}_x^{-1}\mathbf{k} \tag{19}$$

This quantity collapses the entire vector of correlations, the inter-signal dependence structure $\mathbf{R}_x^{-1}$, and the evaluation point $\mathbf{k}$ into a single number, so that criterion (18) simplifies and reads:

$$|\mu_r + \sigma_r\eta| > c \tag{20}$$

The Parameter $\mathbf{k}$

The vector $\mathbf{k} \in \mathbb{R}^n$ controls our evaluation point $\mathbf{x} = \boldsymbol{\mu}_x + \mathbf{D}_x\mathbf{k}$ and has a probabilistic interpretation. Intuitively, one might set all elements in $\mathbf{k}$ to $+3$, i.e. three standard deviations, and test if the resulting expected return is profitable. However, this isolated approach breaks the probabilistic guarantees of our evaluation this parameter should encode. Instead, we must take into account the multivariate signal distribution, which is captured by the squared *Mahalanobis* distance κ^2:

$$\kappa^2 = (\mathbf{x} - \boldsymbol{\mu}_x)^\top \boldsymbol{\Sigma}_x^{-1} (\mathbf{x} - \boldsymbol{\mu}_x) \tag{21}$$

$$= (\mathbf{D}_x\mathbf{k})^\top (\mathbf{D}_x\mathbf{R}_x\mathbf{D}_x)^{-1} (\mathbf{D}_x\mathbf{k}) \tag{22}$$

$$= \mathbf{k}^\top \mathbf{D}_x\mathbf{D}_x^{-1}\mathbf{R}_x^{-1}\mathbf{D}_x^{-1}\mathbf{D}_x\mathbf{k} \tag{23}$$

$$= \mathbf{k}^\top \mathbf{R}_x^{-1}\mathbf{k} \tag{24}$$

However, setting κ is still not sufficient. If we were to test an arbitrary point $\mathbf{k}$ that lies on the boundary of an ellipsoid with distance κ, and this point fails the profitability criterion, we cannot mathematically guarantee that the *entire* ellipsoid is unprofitable. This is because the expected return is a linear hyperplane, it might intersect the ellipsoid such that other, equally likely joint signal realizations on the same κ-ellipsoid yield more profitable returns. Therefore, to establish a strict lower bound on the fraction of unprofitable realizations, we must fix our "*probability budget*" $\kappa > 0$ upfront and subsequently find the most profitable point on that specific κ-ellipsoid. If the correlation $\boldsymbol{\rho}$ fails to exceed the cost threshold c even at this optimum, i.e., by (16) $\mu_r + \boldsymbol{\beta}^\top \mathbf{D}_x\mathbf{k}^* \leq c$ (for a long position) or $\mu_r + \boldsymbol{\beta}^\top \mathbf{D}_x\mathbf{k}^* \geq -c$ (for a short position), then by linearity, the entire ellipsoid and its tangent half-space extending in the adverse direction must be unprofitable. Since the variable component here is $\boldsymbol{\beta}^\top \mathbf{D}_x\mathbf{k}$, the optimization problems for finding the exact evaluation points that extremize it for both long (maximum) and short (minimum) are:

$$\max_{\mathbf{k}} / \min_{\mathbf{k}} \quad \boldsymbol{\beta}^\top \mathbf{D}_x\mathbf{k} \quad \text{s.t.} \quad \mathbf{k}^\top \mathbf{R}_x^{-1}\mathbf{k} = \kappa^2 \tag{25}$$

We can solve this analytically using *Lagrange* multipliers, which naturally yields both the global maximum and minimum simultaneously. We define the *Lagrangian* $\mathcal{L}(\mathbf{k}, \lambda)$ and take the first-order condition with respect to $\mathbf{k}$:

$$\mathcal{L}(\mathbf{k},\lambda) = \boldsymbol{\beta}^\top \mathbf{D}_x \mathbf{k} - \lambda(\mathbf{k}^\top \mathbf{R}_x^{-1}\mathbf{k} - \kappa^2) \tag{26}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{k}} = \mathbf{D}_x \boldsymbol{\beta} - 2\lambda \mathbf{R}_x^{-1}\mathbf{k} = \mathbf{0} \tag{27}$$

$$\mathbf{k} = \frac{1}{2\lambda}\mathbf{R}_x \mathbf{D}_x \boldsymbol{\beta} \tag{28}$$

In (27) we set the gradient to zero, and in (28) we solve for $\mathbf{k}$. To find the multiplier λ, we substitute (28) back into the constraint $\mathbf{k}^\top \mathbf{R}_x^{-1}\mathbf{k} = \kappa^2$:

$$\left(\frac{1}{2\lambda}\mathbf{R}_x \mathbf{D}_x \boldsymbol{\beta}\right)^\top \mathbf{R}_x^{-1} \left(\frac{1}{2\lambda}\mathbf{R}_x \mathbf{D}_x \boldsymbol{\beta}\right) = \kappa^2 \tag{29}$$

$$\frac{1}{4\lambda^2}\boldsymbol{\beta}^\top \mathbf{D}_x \mathbf{R}_x \mathbf{R}_x^{-1} \mathbf{R}_x \mathbf{D}_x \boldsymbol{\beta} = \kappa^2 \tag{30}$$

$$\frac{1}{4\lambda^2}\boldsymbol{\beta}^\top \boldsymbol{\Sigma}_x \boldsymbol{\beta} = \kappa^2 \tag{31}$$

$$\frac{1}{2\lambda} = \pm \frac{\kappa}{\sqrt{\boldsymbol{\beta}^\top \boldsymbol{\Sigma}_x \boldsymbol{\beta}}} \tag{32}$$

In (30) we expand the transpose and group the terms. In (31) we cancel $\mathbf{R}_x \mathbf{R}_x^{-1} = \mathbf{I}$ and recognize the covariance matrix $\boldsymbol{\Sigma}_x = \mathbf{D}_x \mathbf{R}_x \mathbf{D}_x$. In (32) we isolate the scaling factor $1/(2\lambda)$. Because taking the square root yields a $\pm$ solution, the *Lagrange* method perfectly captures both extremes: the positive root corresponds to the maximum (long), and the negative root corresponds to the minimum (short). Substituting this back into (28) yields the closed-form analytical solution for the optimal evaluation points:

$$\mathbf{k}^* = \pm\kappa \frac{\mathbf{R}_x \mathbf{D}_x \boldsymbol{\beta}}{\sqrt{\boldsymbol{\beta}^\top \boldsymbol{\Sigma}_x \boldsymbol{\beta}}} \tag{33}$$

$$= \tilde{k}\frac{\mathbf{R}_x \mathbf{D}_x \boldsymbol{\beta}}{\sigma_Y} \tag{34}$$

where we define $\tilde{k} \in \{-\kappa, \kappa\}$ to absorb the sign depending on the trade direction. Because the conditional expected return is a linear combination of the signals, the optimization above is equivalent to projecting the entire n-dimensional signal space onto a single 1-dimensional scalar axis $Y = \boldsymbol{\beta}^\top \mathbf{x}$, with mean $\mu_Y = \boldsymbol{\beta}^\top \boldsymbol{\mu}_x$ and variance $\sigma_Y^2 = \boldsymbol{\beta}^\top \boldsymbol{\Sigma}_x \boldsymbol{\beta}$. At the optimum $\mathbf{k}^*$, the deviation from the mean along this axis reaches its extremum. By operating on this 1-dimensional projection, we can bound the probability of the unprofitable tangent half-space using the one-sided *Cantelli* inequality:

If $\tilde{k} = \kappa > 0$, the optimal evaluation point lies above the mean. All realizations satisfying $Y \leq \mu_Y + \kappa\sigma_Y$ are closer to (or on the opposite side of) the mean and therefore fail. By the *Cantelli* inequality with $\kappa\sigma_Y > 0$:

$$P(Y - \mu_Y \geq \kappa\sigma_Y) \leq \frac{\sigma_Y^2}{\sigma_Y^2 + (\kappa\sigma_Y)^2} \tag{35}$$

$$= \frac{\sigma_Y^2}{\sigma_Y^2(1+\kappa^2)} = \frac{1}{1+\kappa^2} \tag{36}$$

$$P(Y \leq \mu_Y + \kappa\sigma_Y) \geq 1 - \frac{1}{1+\kappa^2} = \frac{\kappa^2}{1+\kappa^2} \tag{37}$$

In (35) we apply the *Cantelli* inequality, $P(Y - \mu_Y \geq d) \leq \sigma_Y^2/(\sigma_Y^2 + d^2)$ for any distance $d > 0$, setting $d = \kappa\sigma_Y$, which is strictly positive. In (36) we factor σ_Y^2 from the denominator and cancel. In (37) we take the complement: since $P(Y - \mu_Y > \kappa\sigma_Y) \leq P(Y - \mu_Y \geq \kappa\sigma_Y) \leq 1/(1+\kappa^2)$, it follows that $P(Y \leq \mu_Y + \kappa\sigma_Y) = 1 - P(Y - \mu_Y > \kappa\sigma_Y) \geq \kappa^2/(1+\kappa^2)$.

If $\tilde{k} = -\kappa < 0$, the optimal evaluation point lies below the mean. All realizations satisfying $Y \geq \mu_Y - \kappa\sigma_Y$ are closer to (or on the opposite side of) the mean and therefore fail. Applying the *Cantelli* inequality with $-\kappa\sigma_Y < 0$:

$$P(Y - \mu_Y \geq -\kappa\sigma_Y) \geq \frac{(-\kappa\sigma_Y)^2}{(-\kappa\sigma_Y)^2 + \sigma_Y^2} \tag{38}$$

$$= \frac{\kappa^2\sigma_Y^2}{\sigma_Y^2(\kappa^2+1)} = \frac{\kappa^2}{1+\kappa^2} \tag{39}$$

$$P(Y \geq \mu_Y - \kappa\sigma_Y) \geq \frac{\kappa^2}{1+\kappa^2} \tag{40}$$

In (38) we apply the *Cantelli* inequality, $P(Y - \mu_Y \geq d) \geq d^2/(d^2 + \sigma_Y^2)$ for any distance $d < 0$, setting $d = -\kappa\sigma_Y$. In (39) we expand the square, factor σ_Y^2 from the denominator, and cancel it with the numerator. In (40) we rearrange the inequality inside the probability measure to isolate Y. Since this form directly bounds the correct side, we bypass the need to calculate the complement.

In both cases, if the combined signal strength $\boldsymbol{\rho}$ fails to clear c at the boundary controlled by κ, the signals are economically non-viable for at least a fraction $\kappa^2/(1+\kappa^2)$ of realizations, which might be too large. A smaller κ raises the bar on $\boldsymbol{\rho}$ because a lower fraction of unprofitable realizations is accepted, whereas a larger κ lowers the bar because a higher fraction of unprofitable realizations is accepted.

Case Distinction

The absolute value in (20) splits into two cases, depending on whether the expression inside is strictly positive or strictly negative:

$$\text{(A)} \quad \mu_r + \sigma_r \eta > c \tag{41}$$

$$\text{(B)} \quad \mu_r + \sigma_r \eta < -c \tag{42}$$

Case (A) corresponds to the signals pushing expected returns above the positive threshold $+c$ (profitable for a long position), while Case (B) corresponds to pushing expected returns below $-c$ (profitable for a short position). Both can be checked independently, and a set of signals may satisfy one, both, or neither.

Case (A): Long Profitability

We rearrange (41) by moving μ_r to the right-hand side, dividing by $\sigma_r > 0$, and expanding η from (19) evaluated at the optimum $\mathbf{k}^*_{max}$:

$$\sigma_r \eta > c - \mu_r \tag{43}$$

$$\eta > \frac{c - \mu_r}{\sigma_r} \tag{44}$$

$$\boldsymbol{\rho}^\top \mathbf{R}_x^{-1} \mathbf{k}^*_{max} > \frac{c - \mu_r}{\sigma_r} \tag{45}$$

$$\boldsymbol{\rho}^\top \mathbf{R}_x^{-1} \left(\kappa \frac{\mathbf{R}_x \mathbf{D}_x \boldsymbol{\beta}}{\sqrt{\boldsymbol{\beta}^\top \boldsymbol{\Sigma}_x \boldsymbol{\beta}}} \right) > \frac{c - \mu_r}{\sigma_r} \tag{46}$$

$$\kappa \frac{\boldsymbol{\rho}^\top \mathbf{D}_x \boldsymbol{\beta}}{\sqrt{\boldsymbol{\beta}^\top \boldsymbol{\Sigma}_x \boldsymbol{\beta}}} > \frac{c - \mu_r}{\sigma_r} \tag{47}$$

$$\kappa \frac{\boldsymbol{\rho}^\top \mathbf{D}_x (\sigma_r \mathbf{D}_x^{-1} \mathbf{R}_x^{-1} \boldsymbol{\rho})}{\sqrt{(\sigma_r \mathbf{D}_x^{-1} \mathbf{R}_x^{-1} \boldsymbol{\rho})^\top \boldsymbol{\Sigma}_x (\sigma_r \mathbf{D}_x^{-1} \mathbf{R}_x^{-1} \boldsymbol{\rho})}} > \frac{c - \mu_r}{\sigma_r} \tag{48}$$

$$\kappa \frac{\sigma_r \boldsymbol{\rho}^\top \mathbf{R}_x^{-1} \boldsymbol{\rho}}{\sqrt{\sigma_r^2 \boldsymbol{\rho}^\top \mathbf{R}_x^{-1} \boldsymbol{\rho}}} > \frac{c - \mu_r}{\sigma_r} \tag{49}$$

$$\kappa \sqrt{\boldsymbol{\rho}^\top \mathbf{R}_x^{-1} \boldsymbol{\rho}} > \frac{c - \mu_r}{\sigma_r} \tag{50}$$

In (43) we move μ_r to the right-hand side. In (44) we divide by $\sigma_r > 0$, which preserves the inequality direction. In (45) we expand η by (19) and evaluate it at the optimum $\mathbf{k}^*_{max}$ for a long position according to (33). In (46) we substitute $\mathbf{k}^*_{max}$ using the positive root in (33). In (47) we cancel $\mathbf{R}_x^{-1}\mathbf{R}_x = \mathbf{I}$. In (48) we substitute $\boldsymbol{\beta} = \sigma_r \mathbf{D}_x^{-1} \mathbf{R}_x^{-1} \boldsymbol{\rho}$ by (12) simultaneously into the numerator and the denominator. In (49) we cancel $\mathbf{D}_x\mathbf{D}_x^{-1} = \mathbf{I}$ in the numerator; in the denominator we expand $\mathbf{\Sigma}_x = \mathbf{D}_x\mathbf{R}_x\mathbf{D}_x$, cancel the adjacent inverse pairs $\mathbf{D}_x^{-1}\mathbf{D}_x$ and $\mathbf{R}_x^{-1}\mathbf{R}_x$, and pull out the scalar σ_r^2. Finally, in (50) we pull σ_r out of the square root and cancel it with the numerator.

The admissible region for $\boldsymbol{\rho}$ is the exterior of a ball centered at the origin in the *Mahalanobis* space defined by $\mathbf{R}_x^{-1}$, with squared radius $(c-\mu_r)^2/(\kappa^2\sigma_r^2)$. Notably, if $\mu_r > c$, the right-hand side in (50) turns negative, and the criterion is automatically satisfied since the unconditional expected return already exceeds the cost threshold c.

Case (B): Short Profitability

We rearrange (42) analogously to Case (A). Moving μ_r to the right-hand side, dividing by $\sigma_r > 0$, expanding η from (19), and evaluating at the optimum $\mathbf{k}^*_{min}$ for a short position using the negative root in (33):

$$\boldsymbol{\rho}^\top\mathbf{R}_x^{-1}\mathbf{k}^*_{min} < \frac{-(c+\mu_r)}{\sigma_r} \tag{51}$$

$$-\kappa\sqrt{\boldsymbol{\rho}^\top\mathbf{R}_x^{-1}\boldsymbol{\rho}} < \frac{-(c+\mu_r)}{\sigma_r} \tag{52}$$

We follow the same algebraic steps as in (46)–(50), substituting $\mathbf{k}^*_{min}$, replacing $\boldsymbol{\beta}$ by (12), and simplifying. The only difference is the negative sign from $\mathbf{k}^*_{min} = -\kappa\mathbf{R}_x\mathbf{D}_x\boldsymbol{\beta}/\sqrt{\boldsymbol{\beta}^\top\mathbf{\Sigma}_x\boldsymbol{\beta}}$. Multiplying (52) by -1:

$$\kappa\sqrt{\boldsymbol{\rho}^\top\mathbf{R}_x^{-1}\boldsymbol{\rho}} > \frac{c+\mu_r}{\sigma_r} \tag{53}$$

Analogously, if $\mu_r < -c$, the right-hand side in (53) turns negative, and the criterion is automatically satisfied since the unconditional expected return already lies below the negative cost threshold $-c$.

Application

Given a set of n signals with unconditional expected return μ_r, a signal-return correlation vector $\boldsymbol{\rho}$, signal correlation matrix $\mathbf{R}_x$, return volatility σ_r, and a cost threshold $c \in \mathbb{R}_{>0}$, the procedure is as follows:

First, choose the *Mahalanobis* distance κ according to how selective you wish to be, noting that it determines the minimum fraction $\kappa^2/(1+\kappa^2)$ of joint signal realizations that is unprofitable by the *Cantelli* inequality. Second, compute $\kappa\sqrt{\boldsymbol{\rho}^\top\mathbf{R}_x^{-1}\boldsymbol{\rho}} = \kappa\sqrt{R^2}$. Third, check if this exceeds $(c-\mu_r)/\sigma_r$ for long profitability (50) and/or $(c+\mu_r)/\sigma_r$ for short profitability (53), keeping in mind that both cases can be checked independently.

Acknowledgement

Many thanks to istwine, who took the time to carefully read through this note and pointed out important conceptual mistakes.