The Use of Kernels in the Portfolio Optimization Problem

Thierry Bazier-Matte

January 26, 2017

Contents

1 Original Problem **1**
1.1 Prescribed investment 3

2 Alternate problem **3**
2.1 Prescribed investment 5

1 Original Problem

Let us consider the following problem, optimized over $q \in \mathscr{R}^p$:

$$\text{minimiser} \quad \sum_{i=1}^{n} \ell(r_i\, q^T x_i) + n\lambda \|q\|^2, \tag{1}$$

where $\ell = -u$. Alternatively, this problem can be respecified using slack vector $\xi \in \mathscr{R}^n$ as

$$\begin{aligned} \text{minimiser} \quad & \sum_{i=1}^{n} \ell(\xi_i) + n\lambda \|q\|^2 \\ \text{tel que} \quad & \xi_i = r_i\, q^T x_i. \end{aligned} \tag{2}$$

Let $\alpha \in \mathscr{R}^n$. The Lagrangian of (2) can be written as

$$\mathscr{L}(q, \xi, \alpha) = \sum_{i=1}^{n} \ell(\xi_i) + n\lambda \|q\|^2 + \sum_{i=1}^{n} \alpha_i (r_i q^T x_i - \xi_i). \tag{3}$$

Because the objective (2) is convex and its constraints are affine in q and ξ, Slater's theorem states that the duality gap of the problem is zero. In other words, solving (1) is equivalent to maximizing the Lagrange dual function g over α:

$$\text{maximiser} \quad g(\alpha) = \inf_{q,\xi} \mathscr{L}(q, \xi, \alpha). \tag{4}$$

Now, note that

$$
\begin{aligned}
g(\alpha) &= \inf_{q,\xi}\left\{\sum_{i=1}^{n}\ell(\xi_i) + n\lambda\|q\|^2 + \sum_{i=1}^{n}\alpha_i(r_i q^T x_i - \xi_i)\right\} && (5)\\
&= \inf_{\xi}\left\{\sum_{i=1}^{n}\ell(\xi_i) - \alpha^T\xi\right\} + \inf_{q}\left\{\sum_{i=1}^{n}\alpha_i r_i q^T x_i + n\lambda\|q\|^2\right\} && (6)\\
&= -\sup_{\xi}\left\{a^T\xi - \sum_{i=1}^{n}\ell(\xi_i)\right\} + \inf_{q}\left\{\sum_{i=1}^{n}\alpha_i r_i q^T x_i + n\lambda\|q\|^2\right\} && (7)\\
&= -\sum_{i=1}^{n}\ell^*(\alpha_i) + \inf_{q}\left\{\sum_{i=1}^{n}\alpha_i r_i q^T x_i + n\lambda\|q\|^2\right\}. && (8)
\end{aligned}
$$

Where ℓ^* is the convex conjugate of the loss function and is defined by

$$\ell(\alpha_i) = \sup_{\xi_i}\left\{\alpha_i\xi_i - \ell(\xi_i)\right\}. \tag{9}$$

Note that the identity

$$f(\xi_1,\ldots,\xi_n) = \sum_{i=1}^{n}\ell(\xi_i) \Longrightarrow f^*(\xi_1,\ldots,\xi_n) = \sum_{i=1}^{n}\ell^*(\xi_i) \tag{10}$$

was used. Consider now the second part of (8). Since the expression is differentiable, we can analytically solve for q:

$$\nabla_q\left\{\sum_{i=1}^{n}\alpha_i r_i q^T x_i + n\lambda\|q\|^2\right\} = 0 \tag{11}$$

implies that

$$q = -\frac{1}{2n\lambda}\sum_{i=1}^{n}\alpha_i r_i x_i \tag{12}$$

at the infimum.

Using (12), we can eliminate q from (8), so that

$$
\begin{aligned}
g(\alpha) &= -\sum_{i=1}^{n}\ell^*(\alpha_i) - \frac{1}{2n\lambda}\sum_{i,j=1}^{n}\alpha_i\alpha_j r_i r_j x_i^T x_j + \frac{1}{4n\lambda}\sum_{i,j=1}^{n}\alpha_i\alpha_j r_i r_j x_i^T x_j && (13)\\
&= -\sum_{i=1}^{n}\ell^*(\alpha_i) - \frac{1}{4n\lambda}(\alpha\circ r)^T K(\alpha\circ r). && (14)
\end{aligned}
$$

Therefore, in its dual form, the problem (1) is equivalent to solving

$$\text{minimiser}\quad \sum_{i=1}^{n}\ell^*(\alpha_i) + \frac{1}{4n\lambda}(\alpha\circ r)^T K(\alpha\circ r). \tag{15}$$

1.1 Prescribed investment

In its original form, given a feature vector $\tilde{x}$, the algorithm (1) suggests an investment size of $p_0 = q^T \tilde{x}$, where q is the trained value obtained by optimizing (1). In the dual formulation (15), with optimal value α, we have from (12):

$$
\begin{aligned}
p_0 &= q^T x_0 && (16)\\
&= -\frac{1}{2n\lambda} \sum_{i=1}^{n} \alpha_i r_i x_i^T x_0. && (17)
\end{aligned}
$$

[Todo: Insert kernel formulation with vector ϕ.]

2 Alternate problem

We now consider a new problem, slightly different from (1) where a regularization based on the sum of the square of the investment sizes $q^T x_i$ is applied:

$$
\text{minimiser} \quad \sum_{i=1}^{n} \ell(r_i\, q^T x_i) + \gamma \sum_{i=1}^{n} (q^T x_i)^2 + n\lambda \|q\|^2. \tag{18}
$$

Again, this problem can be respecified using slack vector $\xi \in \mathscr{R}^n$ as

$$
\begin{aligned}
\text{minimiser} \quad & \sum_{i=1}^{n} \ell(\xi_i) + \gamma \sum_{i=1}^{n} (\xi_i / r_i)^2 + n\lambda \|q\|^2 \\
\text{tel que} \quad & \xi_i = r_i\, q^T x_i.
\end{aligned} \tag{19}
$$

The constraints in (19) are again affine, so that Slater's theorem apply.

The lagrangian of (19) is

$$
\mathscr{L}(q, \xi, \alpha) = \sum_{i=1}^{n} \ell(\xi_i) + \gamma \sum_{i=1}^{n} (\xi_i / r_i)^2 + n\lambda \|q\|^2 + \sum_{i=1}^{n} \alpha_i (r_i q^T x_i - \xi_i), \tag{20}
$$

and we seek its infimum over (q, ξ).

$$
\begin{aligned}
&\inf_{q,\xi} \left\{ \sum_{i=1}^{n} \ell(\xi_i) + \gamma \sum_{i=1}^{n} (\xi_i / r_i)^2 + n\lambda \|q\|^2 + \sum_{i=1}^{n} \alpha_i (r_i q^T x_i - \xi_i) \right\} && (21)\\
&\quad = \inf_{\xi} \left\{ \sum_{i=1}^{n} \ell(\xi_i) + \gamma \sum_{i=1}^{n} (\xi_i / r_i)^2 - \alpha^T \xi \right\} + \inf_{q} \left\{ \sum_{i=1}^{n} \alpha_i r_i q^T x_i - n\lambda \|q\|^2 \right\} && (22)\\
&\quad = -\sup_{\xi} \left\{ a^T \xi - \left(\sum_{i=1}^{n} \ell(\xi_i) + \gamma \sum_{i=1}^{n} (\xi_i / r_i)^2 \right) \right\} - \frac{1}{4n\lambda} (\alpha \circ r)^T K (\alpha \circ r). && (23)
\end{aligned}
$$

Let $f_i(\xi_i) := h_1(\xi_i)+h_2(\xi_i) = \ell(\xi_i)+\gamma(\xi_i/r_i)^2$. Then, using (10), the first expression of (23) can be restated as

$$-\sup_{\xi}\left\{\alpha^T\xi - \sum_{i=1}^{n} f_i(\xi_i)\right\} = -\sum_{i=1}^{n} f_i^*(\alpha_i). \tag{24}$$

Let us introduce another identity:

$$(h_1+h_2)^*(\alpha_i) = \inf_{\alpha_i'+\alpha_i''=\alpha_i}\{h_1^*(\alpha_i') + h_2^*(\alpha_i'')\}. \tag{25}$$

Using (25), (24) can be written as

$$-\sum_{i=1}^{n} f_i^*(\alpha_i) = -\sum_{i=1}^{n}(h_1+h_2)^*(\xi_i) \tag{26}$$

$$= -\sum_{i=1}^{n}\inf_{\alpha_i'+\alpha_i''=\alpha_i}\{h_1^*(\alpha_i') + h_2^*(\alpha_i'')\}. \tag{27}$$

The first conjugate function h_1^* is simply ℓ^*. The second conjugate function can be derived analytically:

$$h_2^*(\alpha_i'') = \sup_{\xi_i}\{\alpha_i''\xi_i - h_2(\xi_i)\} \tag{28}$$

$$= \sup_{\xi_i}\{\alpha_i''\xi_i - \gamma(\xi_i/r_i)^2\}. \tag{29}$$

The supremum occurs when

$$\xi_i = \frac{r_i^2}{2\gamma}\alpha_i''. \tag{30}$$

Therefore, (29) simplifies to

$$h_2^*(\alpha_i'') = \frac{r_i^2}{4\gamma}(\alpha_i'')^2. \tag{31}$$

Putting it all back together, the dual of (18) is

$$-\sum_{i=1}^{n}\inf_{\alpha_i'+\alpha_i''=\alpha_i}\left\{\ell^*(\alpha_i') + \frac{r_i^2}{4\gamma}(\alpha_i'')^2\right\} - \frac{1}{4n\lambda}(\alpha\circ r)^T K(\alpha\circ r), \tag{32}$$

which is equivalent to

$$-\sum_{i=1}^{n}\ell^*(\alpha_i) - \frac{1}{4\gamma}\sum_{i=1}^{n}(r_i\,\beta_i)^2 - \frac{1}{4n\lambda}(r\circ(\alpha+\beta))^T K(r\circ(\alpha+\beta)), \tag{33}$$

with new optimization variables $\alpha = \alpha', \beta = \alpha'' \in \mathscr{R}^n$. The dual optimization problem is therefore

$$\text{minimiser}\quad \sum_{i=1}^{n}\ell^*(\alpha_i) + \frac{1}{4\gamma}\|r\circ\beta\|^2 + \frac{1}{4n\lambda}(r\circ(\alpha+\beta))^T K(r\circ(\alpha+\beta)). \tag{34}$$

2.1 Prescribed investment

[Todo:]