The Use of Kernels in the Portfolio Optimization Problem

Thierry Bazier-Matte

February 3, 2017

1 Introduction aux fonctions de décisions non linéaires

1.1 Introduction

Soit $\kappa : \boldsymbol{X} \times \boldsymbol{X} \to \mathscr{R}$ un noyau semi-défini positif, $\boldsymbol{H}$ l'espace de Hilbert à noyau reproduisant induit par κ et $K \in \mathscr{R}^{n \times n}$ la matrice associée à l'ensemble d'échantillonage $S_n \sim M^n$. Le problème d'optimisation de portefeuille régularisé s'exprime alors par

$$\underset{q \in \boldsymbol{H}}{\text{maximiser}} \quad n^{-1} \sum_{i=1}^{n} u(r_i\, q(x_i)) - \lambda \|q\|_{\boldsymbol{H}}^2. \tag{1}$$

Tel que mentionné, la dimension de $\boldsymbol{H}$ est possiblement infinie, ce qui rend numériquement impossible la recherche d'une solution $q^\star$. Toutefois, le théorème de la représentation permet de rendre le problème résoluble.

Théorème. *Toute solution $q^\star$ de* (1) *repose dans le sous-espace vectoriel engendré par l'ensemble des n fonctions $\{\phi_i\}$, où $\phi_i = \kappa(x_i, \cdot)$. Numériquement, il existe un vecteur $\alpha^\star \in \mathscr{R}^n$ tel que,*

$$q^\star = \sum_{i=1}^{n} \alpha_i^\star \phi_i = (\alpha^\star)^T \phi. \tag{2}$$

Démonstration. Voir [MRT12], Théorème 5.4 pour une démonstration tenant compte d'un objectif régularisé général. La démonstration est dûe à [KW71]. □

Le théorème de la représentation permet donc de chercher une solution dans un espace à n dimensions, plutôt que la dimension possiblement infinie de $\boldsymbol{H}$. En effet, puisque

$$q^\star = \sum_{i=1}^{n} \alpha_i \phi_i, \tag{3}$$

où $\alpha \in \mathscr{R}^n$ **[Todo:** Espace cotangant???**]**, on peut donc restreindre le domaine d'optimisation à $\mathscr{R}^n$. L'objectif de (1) devient alors

$$n^{-1} \sum_{i=1}^{n} u(r_i \textstyle\sum_{j=1}^{n} \alpha_j \phi_j(x_i)) - \lambda \langle q, q \rangle_{\boldsymbol{H}}. \tag{4}$$

Le premier terme se réexprime comme

$$n^{-1} \sum_{i=1}^{n} u(r_i \, \alpha^T \phi(x_i)), \tag{5}$$

alors qu'en employant les propriétés de linéarité du produit intérieur, on transforme le second terme par

$$\begin{aligned}
\langle q, q \rangle_{\boldsymbol{H}}^2 &= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \langle \phi_i, \phi_j \rangle_{\boldsymbol{H}} &(6)\\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \kappa(x_i, x_j) &(7)\\
&= \alpha^T K \alpha. &(8)
\end{aligned}$$

De sorte que le problème général (1) peut se reformuler par

$$\boxed{\underset{\alpha \in \mathscr{R}^n}{\text{maximiser}} \quad n^{-1} \sum_{i=1}^{n} u(r_i \alpha^T \phi(x_i)) - \lambda \, \alpha^T K \alpha} \quad . \tag{9}$$

1.2 Dualité

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

$$\text{minimiser} \quad \sum_{i=1}^{n} \ell(r_i \, q^T x_i) + n\lambda \|q\|^2, \tag{10}$$

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

$$\begin{aligned}
\text{minimiser} \quad & \sum_{i=1}^{n} \ell(\xi_i) + n\lambda \|q\|^2 \\
\text{tel que} \quad & \xi_i = r_i \, q^T x_i.
\end{aligned} \tag{11}$$

Let $\alpha \in \mathscr{R}^n$. The Lagrangian of (11) can be written as

$$\mathscr{L}(q, \xi, \alpha) = \sum_{i=1}^{n} \ell(\xi_i) + n\lambda \|q\|^2 + \sum_{i=1}^{n} \alpha_i (r_i q^T x_i - \xi_i). \tag{12}$$

Because the objective (11) 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 (10) is equivalent to maximizing the Lagrange dual function g over α:

$$\text{maximiser} \quad g(\alpha) = \inf_{q,\xi} \mathscr{L}(q, \xi, \alpha). \tag{13}$$

Now, note that

$$\begin{aligned} g(\alpha) &= \inf_{q,\xi} \left\{ \sum_{i=1}^{n} \ell(\xi_i) + n\lambda \|q\|^2 + \sum_{i=1}^{n} \alpha_i (r_i q^T x_i - \xi_i) \right\} && (14) \\ &= \inf_{\xi} \left\{ \sum_{i=1}^{n} \ell(\xi_i) - \alpha^T \xi \right\} + \inf_{q} \left\{ \sum_{i=1}^{n} \alpha_i r_i q^T x_i + n\lambda \|q\|^2 \right\} && (15) \\ &= -\sup_{\xi} \left\{ a^T \xi - \sum_{i=1}^{n} \ell(\xi_i) \right\} + \inf_{q} \left\{ \sum_{i=1}^{n} \alpha_i r_i q^T x_i + n\lambda \|q\|^2 \right\} && (16) \\ &= -\sum_{i=1}^{n} \ell^*(\alpha_i) + \inf_{q} \left\{ \sum_{i=1}^{n} \alpha_i r_i q^T x_i + n\lambda \|q\|^2 \right\}. && (17) \end{aligned}$$

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

$$\ell(\alpha_i) = \sup_{\xi_i} \left\{ \alpha_i \xi_i - \ell(\xi_i) \right\}. \tag{18}$$

Note that the identity

$$f(\xi_1, \ldots, \xi_n) = \sum_{i=1}^{n} \ell(\xi_i) \Longrightarrow f^*(\xi_1, \ldots, \xi_n) = \sum_{i=1}^{n} \ell^*(\xi_i) \tag{19}$$

was used. Consider now the second part of (17). Since the expression is differentiable, we can analytically solve for q:

$$\nabla_q \left\{ \sum_{i=1}^{n} \alpha_i r_i q^T x_i + n\lambda \|q\|^2 \right\} = 0 \tag{20}$$

implies that

$$q = -\frac{1}{2n\lambda} \sum_{i=1}^{n} \alpha_i r_i x_i \tag{21}$$

at the infimum.

Using (21), we can eliminate q from (17), so that

$$\begin{aligned} g(\alpha) &= -\sum_{i=1}^{n} \ell^*(\alpha_i) - \frac{1}{2n\lambda} \sum_{i,j=1}^{n} \alpha_i \alpha_j r_i r_j x_i^T x_j + \frac{1}{4n\lambda} \sum_{i,j=1}^{n} \alpha_i \alpha_j r_i r_j x_i^T x_j && (22) \\ &= -\sum_{i=1}^{n} \ell^*(\alpha_i) - \frac{1}{4n\lambda} (\alpha \circ r)^T K (\alpha \circ r). && (23) \end{aligned}$$

Therefore, in its dual form, the problem (10) is equivalent to solving

$$\text{minimiser} \quad \sum_{i=1}^{n} \ell^*(\alpha_i) + \frac{1}{4n\lambda}(\alpha \circ r)^T K(\alpha \circ r). \tag{24}$$

Prescribed investment In its original form, given a feature vector $\tilde{x}$, the algorithm (10) suggests an investment size of $p_0 = q^T\tilde{x}$, where q is the trained value obtained by optimizing (10). In the dual formulation (24), with optimal value α, we have from (21):

$$p_0 = q^T x_0 \tag{25}$$

$$= -\frac{1}{2n\lambda}\sum_{i=1}^{n} \alpha_i r_i x_i^T x_0. \tag{26}$$

[Todo: Insert kernel formulation with vector ϕ.**]**

1.3 Alternate problem

We now consider a new problem, slightly different from (10) where a regularization based on the sum of the square of the investment sizes $q^T x_i$ is applied:

$$\text{minimiser} \quad \sum_{i=1}^{n} \ell(r_i\, q^T x_i) + \gamma \sum_{i=1}^{n} (q^T x_i)^2 + n\lambda \|q\|^2. \tag{27}$$

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

$$\begin{aligned} \text{minimiser} \quad & \sum_{i=1}^{n} \ell(\xi_i) + \gamma \sum_{i=1}^{n} (\xi_i/r_i)^2 + n\lambda \|q\|^2 \\ \text{tel que} \quad & \xi_i = r_i\, q^T x_i. \end{aligned} \tag{28}$$

The constraints in (28) are again affine, so that Slater's theorem apply.

The lagrangian of (28) is

$$\mathscr{L}(q, \xi, \alpha) = \sum_{i=1}^{n} \ell(\xi_i) + \gamma \sum_{i=1}^{n} (\xi_i/r_i)^2 + n\lambda\|q\|^2 + \sum_{i=1}^{n} \alpha_i (r_i q^T x_i - \xi_i), \tag{29}$$

and we seek its infimum over (q, ξ).

$$\inf_{q,\xi} \left\{ \sum_{i=1}^{n} \ell(\xi_i) + \gamma \sum_{i=1}^{n} (\xi_i/r_i)^2 + n\lambda\|q\|^2 + \sum_{i=1}^{n} \alpha_i (r_i q^T x_i - \xi_i) \right\} \tag{30}$$

$$= \inf_{\xi} \left\{ \sum_{i=1}^{n} \ell(\xi_i) + \gamma \sum_{i=1}^{n} (\xi_i/r_i)^2 - \alpha^T \xi \right\} + \inf_{q} \left\{ \sum_{i=1}^{n} \alpha_i r_i q^T x_i - n\lambda\|q\|^2 \right\} \tag{31}$$

$$= -\sup_{\xi} \left\{ a^T \xi - \left(\sum_{i=1}^{n} \ell(\xi_i) + \gamma \sum_{i=1}^{n} (\xi_i / r_i)^2 \right) \right\} - \frac{1}{4n\lambda} (\alpha \circ r)^T K (\alpha \circ r). \tag{32}$$

Let $f_i(\xi_i) := h_1(\xi_i) + h_2(\xi_i) = \ell(\xi_i) + \gamma(\xi_i/r_i)^2$. Then, using (19), the first expression of (32) can be restated as

$$-\sup_{\xi} \left\{ \alpha^T \xi - \sum_{i=1}^{n} f_i(\xi_i) \right\} = -\sum_{i=1}^{n} f_i^*(\alpha_i). \tag{33}$$

Let us introduce another identity:

$$(h_1 + h_2)^*(\alpha_i) = \inf_{\alpha_i' + \alpha_i'' = \alpha_i} \{ h_1^*(\alpha_i') + h_2^*(\alpha_i'') \}. \tag{34}$$

Using (34), (33) can be written as

$$-\sum_{i=1}^{n} f_i^*(\alpha_i) = -\sum_{i=1}^{n} (h_1 + h_2)^*(\xi_i) \tag{35}$$

$$= -\sum_{i=1}^{n} \inf_{\alpha_i' + \alpha_i'' = \alpha_i} \{ h_1^*(\alpha_i') + h_2^*(\alpha_i'') \}. \tag{36}$$

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

$$h_2^*(\alpha_i'') = \sup_{\xi_i} \{ \alpha_i'' \xi_i - h_2(\xi_i) \} \tag{37}$$

$$= \sup_{\xi_i} \{ \alpha_i'' \xi_i - \gamma (\xi_i / r_i)^2 \}. \tag{38}$$

The supremum occurs when

$$\xi_i = \frac{r_i^2}{2\gamma} \alpha_i''. \tag{39}$$

Therefore, (38) simplifies to

$$h_2^*(\alpha_i'') = \frac{r_i^2}{4\gamma} (\alpha_i'')^2. \tag{40}$$

Putting it all back together, the dual of (27) is

$$-\sum_{i=1}^{n} \inf_{\alpha_i' + \alpha_i'' = \alpha_i} \left\{ \ell^*(\alpha_i') + \frac{r_i^2}{4\gamma} (\alpha_i'')^2 \right\} - \frac{1}{4n\lambda} (\alpha \circ r)^T K (\alpha \circ r), \tag{41}$$

which is equivalent to

$$-\sum_{i=1}^{n} \ell^*(\alpha_i) - \frac{1}{4\gamma} \sum_{i=1}^{n} (r_i \, \beta_i)^2 - \frac{1}{4n\lambda} (r \circ (\alpha + \beta))^T K (r \circ (\alpha + \beta)), \tag{42}$$

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

$$\text{minimiser} \quad \sum_{i=1}^{n} \ell^*(\alpha_i) + \frac{1}{4\gamma}\|r \circ \beta\|^2 + \frac{1}{4n\lambda}(r \circ (\alpha + \beta))^T K(r \circ (\alpha + \beta)). \tag{43}$$

Prescribed investment [Todo:]

References

[KW71] George Kimeldorf and Grace Wahba. Some results on tchebycheffian spline functions. *Journal of mathematical analysis and applications*, 33(1):82–95, 1971.

[MRT12] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning*. MIT press, 2012.