Journal of Industrial and Management Optimization

https://www.aimspress.com/journal/jimo

JIMO, 22(6): 2755–2783.
DOI:10.3934/jimo.2026101
Received: 23 February 2026
Revised: 26 April 2026
Accepted: 08 May 2026
Published: 20 May 2026

Research article

Cardinality-constrained maximal predictability portfolios with an ℓ_2 regularization

Katsuhiro Tanaka* **and Rei Yamamoto**

Department of Industrial and Systems Engineering, Keio University, 3-14-1, Hiyoshi, Kohoku-ku, Yokohama-shi, Kanagawa 223-8522, Japan

* **Correspondence:** Email: katsuhiro-tanaka@keio.jp; Tel: +81-45-566-1774.

Appendix

Appendix A Proof of Theorem 1

For convenience, we additionally impose the redundant upper-bound constraints $y_n \leq M$ $(n \in \mathcal{N})$ on Problem (2.12), where M is chosen sufficiently large so that these upper-bound constraints are inactive at an optimal solution.

The Lagrange function of Problem (2.12) is represented as follows:

$$\begin{aligned} &\mathcal{L}(\boldsymbol{y}, \boldsymbol{v}, \boldsymbol{u}; \boldsymbol{\phi}, \boldsymbol{\psi}, b, \boldsymbol{\pi}, \boldsymbol{\xi}, s) \\ := \;& \frac{1}{2\gamma}\boldsymbol{y}^\top \boldsymbol{y} + \boldsymbol{v}^\top \boldsymbol{v} - \boldsymbol{\phi}^\top(\boldsymbol{v} - \boldsymbol{E}\boldsymbol{y}) - \boldsymbol{\psi}^\top(\boldsymbol{u} - \boldsymbol{R}\boldsymbol{y}) \\ &- b(\tilde{\boldsymbol{r}}^\top \boldsymbol{y} - \rho \boldsymbol{1}^\top \boldsymbol{y}) - \boldsymbol{\pi}^\top \boldsymbol{y} - \boldsymbol{\xi}^\top(M\boldsymbol{1} - \boldsymbol{y}) - s(\boldsymbol{u}^\top \boldsymbol{u} - 1), \end{aligned} \tag{A.1}$$

where

$$b \geq 0, \quad \boldsymbol{\pi} \geq \boldsymbol{0}, \quad \boldsymbol{\xi} \geq \boldsymbol{0}, \quad \boldsymbol{\phi} \in \mathbb{R}^T, \quad \boldsymbol{\psi} \in \mathbb{R}^T, \quad s \in \mathbb{R}$$

are Lagrange multipliers and $\boldsymbol{1}$ denotes the all-ones vector of appropriate dimension. Then, the dual problem associated with Problem (2.12) is represented as follows:

$$\max_{\boldsymbol{\phi}, \boldsymbol{\psi}, b, \boldsymbol{\pi}, \boldsymbol{\xi}, s} \; \min_{\boldsymbol{y}, \boldsymbol{v}, \boldsymbol{u}} \; \mathcal{L}(\boldsymbol{y}, \boldsymbol{v}, \boldsymbol{u}; \boldsymbol{\phi}, \boldsymbol{\psi}, b, \boldsymbol{\pi}, \boldsymbol{\xi}, s). \tag{A.2}$$

We focus on the inner minimization problem:

$$\min_{\boldsymbol{y},\boldsymbol{v},\boldsymbol{u}} \mathcal{L}(\boldsymbol{y}, \boldsymbol{v}, \boldsymbol{u}; \boldsymbol{\phi}, \boldsymbol{\psi}, b, \boldsymbol{\pi}, \boldsymbol{\xi}, s). \tag{A.3}$$

Because the objective function of Problem (A.3) is differentiable, the following optimality conditions are satisfied:

$$\nabla_{\boldsymbol{y}} = \frac{1}{\gamma}\boldsymbol{y} - (-\boldsymbol{E}^\top \boldsymbol{\phi} - \boldsymbol{R}^\top \boldsymbol{\psi} + b(\tilde{\boldsymbol{r}} - \rho \mathbf{1}) + \boldsymbol{\pi} - \boldsymbol{\xi}) = 0, \tag{A.4a}$$

$$\nabla_{\boldsymbol{v}} = 2\boldsymbol{v} - \boldsymbol{\phi} = 0, \tag{A.4b}$$

$$\nabla_{\boldsymbol{u}} = -\boldsymbol{\psi} - 2s\boldsymbol{u} = 0. \tag{A.4c}$$

According to (A.4), Problem (A.2) can be represented using the following objective function:

$$-\frac{\gamma}{2}\boldsymbol{w}^\top \boldsymbol{w} - \frac{\boldsymbol{\phi}^\top \boldsymbol{\phi}}{4} + s - M\mathbf{1}^\top \boldsymbol{\xi}, \tag{A.5}$$

where $\boldsymbol{w} \in \mathbb{R}^N$ is defined by

$$\boldsymbol{w} = -\boldsymbol{E}^\top \boldsymbol{\phi} + 2s\boldsymbol{R}^\top \boldsymbol{u} + b(\tilde{\boldsymbol{r}} - \rho \mathbf{1}) + \boldsymbol{\pi} - \boldsymbol{\xi}. \tag{A.6}$$

Then, from (A.4a) and (A.6), we obtain

$$\boldsymbol{y} = \gamma \boldsymbol{w}.$$

As $\gamma > 0$ and $\boldsymbol{y} \geq \mathbf{0}$ in (2.7), it follows that $\boldsymbol{w} \geq \mathbf{0}$.

Moreover, the equality in (A.6) with $\boldsymbol{\pi} \geq \mathbf{0}$ is equivalent to

$$\boldsymbol{w} \geq -\boldsymbol{E}^\top \boldsymbol{\phi} + 2s\boldsymbol{R}^\top \boldsymbol{u} + b(\tilde{\boldsymbol{r}} - \rho \mathbf{1}) - \boldsymbol{\xi}. \tag{A.7}$$

Indeed, if (A.6) holds for some $\boldsymbol{\pi} \geq \mathbf{0}$, then (A.7) immediately follows. Conversely, if (A.7) holds, then one can obtain

$$\boldsymbol{\pi} = \boldsymbol{w} + \boldsymbol{E}^\top \boldsymbol{\phi} - 2s\boldsymbol{R}^\top \boldsymbol{u} - b(\tilde{\boldsymbol{r}} - \rho \mathbf{1}) + \boldsymbol{\xi},$$

which satisfies $\boldsymbol{\pi} \geq \mathbf{0}$ and recovers (A.6). Hence, $\boldsymbol{\pi}$ can be eliminated from the dual formulation.

Thus, Problem (A.2) can be represented as follows:

$$\left|\begin{array}{ll} \max\limits_{\boldsymbol{w},\boldsymbol{\phi},s,b,\boldsymbol{\xi}} \min\limits_{\boldsymbol{u}} & -\frac{\gamma}{2}\boldsymbol{w}^\top \boldsymbol{w} - \frac{\boldsymbol{\phi}^\top \boldsymbol{\phi}}{4} + s - M\mathbf{1}^\top \boldsymbol{\xi} \\ \text{s.t.} & \boldsymbol{w} \geq -\boldsymbol{E}^\top \boldsymbol{\phi} + 2s\boldsymbol{R}^\top \boldsymbol{u} + b(\tilde{\boldsymbol{r}} - \rho \mathbf{1}) - \boldsymbol{\xi}, \\ & \boldsymbol{w} \geq \mathbf{0}, \quad \boldsymbol{\xi} \geq \mathbf{0}, \quad b \geq 0. \end{array}\right. \tag{A.8}$$

By the assumption that M is sufficiently large, the upper-bound constraints $y_n \leq M$ $(n \in \mathcal{N})$ are inactive at an optimal solution of Problem (2.12). Therefore, by the complementary slackness condition,

$$\xi_n(M - y_n) = 0 \qquad (n \in \mathcal{N})$$

and $M - y_n > 0$, which imply

$$\xi_n = 0 \qquad (n \in N).$$

Hence, $\boldsymbol{\xi} = \mathbf{0}$, and Problem (A.8) reduces to

$$\left|\begin{array}{ll} \max\limits_{\boldsymbol{w},\boldsymbol{\phi},s,b} \min\limits_{\boldsymbol{u}} & -\frac{\gamma}{2}\boldsymbol{w}^\top \boldsymbol{w} - \frac{\boldsymbol{\phi}^\top \boldsymbol{\phi}}{4} + s \\ \text{s.t.} & \boldsymbol{w} \geq -\boldsymbol{E}^\top \boldsymbol{\phi} + 2s\boldsymbol{R}^\top \boldsymbol{u} + b(\tilde{\boldsymbol{r}} - \rho \mathbf{1}), \\ & \boldsymbol{w} \geq \mathbf{0}, \quad b \geq 0. \end{array}\right. \tag{A.9}$$

□

Appendix B Proof of Theorem 2

We assume that M is chosen sufficiently large so that the upper-bound constraints $y_n \leq M$ ($n \in \mathcal{N}(z)$) are inactive at an optimal solution of Problem (2.30), as in (2.7).

The Lagrange function of Problem (2.30) is formulated as follows:

$$\begin{aligned}
&\overline{\mathcal{L}}(\boldsymbol{y}_z, \boldsymbol{v}, \boldsymbol{u}; \boldsymbol{\phi}, \boldsymbol{\psi}, b, \boldsymbol{\pi}, \boldsymbol{\xi}, s) \\
:=\ & \frac{1}{2\gamma}\boldsymbol{y}_z^\top \boldsymbol{y}_z + \boldsymbol{v}^\top \boldsymbol{v} - \boldsymbol{\phi}^\top(\boldsymbol{v} - \boldsymbol{E}_z \boldsymbol{y}_z) - \boldsymbol{\psi}^\top(\boldsymbol{u} - \boldsymbol{R}_z \boldsymbol{y}_z) \\
& - b(\tilde{\boldsymbol{r}}_z^\top \boldsymbol{y}_z - \rho \boldsymbol{1}^\top \boldsymbol{y}_z) - \boldsymbol{\pi}^\top \boldsymbol{y}_z - \boldsymbol{\xi}^\top(M\boldsymbol{1} - \boldsymbol{y}_z) - s(\boldsymbol{u}^\top \boldsymbol{u}^*(z) - 1),
\end{aligned} \tag{B.1}$$

where

$$b \geq 0, \quad \boldsymbol{\pi} \geq \boldsymbol{0}, \quad \boldsymbol{\xi} \geq \boldsymbol{0}, \quad \boldsymbol{\phi} \in \mathbb{R}^T, \quad \boldsymbol{\psi} \in \mathbb{R}^T, \quad s \in \mathbb{R}$$

are Lagrange multipliers. Then, the dual problem associated with Problem (2.30) is given by

$$\max_{\boldsymbol{\phi}, \boldsymbol{\psi}, b, \boldsymbol{\pi}, \boldsymbol{\xi}, s} \ \min_{\boldsymbol{y}_z, \boldsymbol{v}, \boldsymbol{u}} \ \overline{\mathcal{L}}(\boldsymbol{y}_z, \boldsymbol{v}, \boldsymbol{u}; \boldsymbol{\phi}, \boldsymbol{\psi}, b, \boldsymbol{\pi}, \boldsymbol{\xi}, s). \tag{B.2}$$

We focus on the inner minimization problem:

$$\min_{\boldsymbol{y}_z, \boldsymbol{v}, \boldsymbol{u}} \ \overline{\mathcal{L}}(\boldsymbol{y}_z, \boldsymbol{v}, \boldsymbol{u}; \boldsymbol{\phi}, \boldsymbol{\psi}, b, \boldsymbol{\pi}, \boldsymbol{\xi}, s). \tag{B.3}$$

Because the objective function of Problem (B.3) is differentiable and convex, the following optimality conditions are satisfied:

$$\nabla_{\boldsymbol{y}_z} = \frac{1}{\gamma}\boldsymbol{y}_z - \left(-\boldsymbol{E}_z^\top \boldsymbol{\phi} - \boldsymbol{R}_z^\top \boldsymbol{\psi} + b(\tilde{\boldsymbol{r}}_z - \rho \boldsymbol{1}) + \boldsymbol{\pi} - \boldsymbol{\xi}\right) = 0, \tag{B.4a}$$

$$\nabla_{\boldsymbol{v}} = 2\boldsymbol{v} - \boldsymbol{\phi} = 0, \tag{B.4b}$$

$$\nabla_{\boldsymbol{u}} = -\boldsymbol{\psi} - 2s\boldsymbol{u}^*(z) = 0. \tag{B.4c}$$

According to (B.4), Problem (B.2) can be reformulated using the following objective function:

$$-\frac{\gamma}{2}\boldsymbol{w}_z^\top \boldsymbol{w}_z - \frac{\boldsymbol{\phi}^\top \boldsymbol{\phi}}{4} + s - M\boldsymbol{1}^\top \boldsymbol{\xi}, \tag{B.5}$$

where

$$\boldsymbol{w}_z = (w_n \mid z_n = 1) \in \mathbb{R}^{|\mathcal{N}(z)|}$$

is defined by

$$\boldsymbol{w}_z = -\boldsymbol{E}_z^\top \boldsymbol{\phi} + 2s\boldsymbol{R}_z^\top \boldsymbol{u}^*(z) + b(\tilde{\boldsymbol{r}}_z - \rho \boldsymbol{1}) + \boldsymbol{\pi} - \boldsymbol{\xi}. \tag{B.6}$$

Then, from (B.4a) and (B.6), we obtain

$$\boldsymbol{y}_z = \gamma \boldsymbol{w}_z.$$

As $\gamma > 0$ and $\boldsymbol{y}_z \geq \boldsymbol{0}$ in (2.28), it follows that $\boldsymbol{w}_z \geq \boldsymbol{0}$.

Moreover, the equality in (B.6) with $\boldsymbol{\pi} \geq \boldsymbol{0}$ is equivalent to

$$\boldsymbol{w}_z \geq -\boldsymbol{E}_z^\top \boldsymbol{\phi} + 2s\boldsymbol{R}_z^\top \boldsymbol{u}^*(z) + b(\tilde{\boldsymbol{r}}_z - \rho \boldsymbol{1}) - \boldsymbol{\xi}. \tag{B.7}$$

Indeed, if (B.6) holds for some $\boldsymbol{\pi} \geq \mathbf{0}$, then (B.7) immediately follows. Conversely, if (B.7) holds, then one can obtain

$$\boldsymbol{\pi} = \boldsymbol{w}_z + \boldsymbol{E}_z^\top \boldsymbol{\phi} - 2s\boldsymbol{R}_z^\top \boldsymbol{u}^*(z) - b(\tilde{\boldsymbol{r}}_z - \rho \mathbf{1}) + \boldsymbol{\xi},$$

which satisfies $\boldsymbol{\pi} \geq \mathbf{0}$ and recovers (B.6). Hence, $\boldsymbol{\pi}$ can be eliminated from the dual formulation.

Thus, Problem (B.2) can be reformulated as

$$\left|\begin{array}{ll} \max\limits_{\boldsymbol{w}_z, \boldsymbol{\phi}, s, b, \boldsymbol{\xi}} & -\dfrac{\gamma}{2}\boldsymbol{w}_z^\top \boldsymbol{w}_z - \dfrac{\boldsymbol{\phi}^\top \boldsymbol{\phi}}{4} + s - M\mathbf{1}^\top \boldsymbol{\xi} \\ \text{s.t.} & \boldsymbol{w}_z \geq -\boldsymbol{E}_z^\top \boldsymbol{\phi} + 2s\boldsymbol{R}_z^\top \boldsymbol{u}^*(z) + b(\tilde{\boldsymbol{r}}_z - \rho \mathbf{1}) - \boldsymbol{\xi}, \\ & \boldsymbol{w}_z \geq \mathbf{0}, \quad \boldsymbol{\xi} \geq \mathbf{0}, \quad b \geq 0. \end{array}\right. \tag{B.8}$$

By the assumption that M is sufficiently large, the upper-bound constraints $y_n \leq M$ ($n \in \mathcal{N}(z)$) are inactive at an optimal solution of Problem (2.30). Therefore, by the complementary slackness condition,

$$\xi_n(M - y_n) = 0 \qquad (n \in \mathcal{N}(z))$$

and $M - y_n > 0$, which imply

$$\xi_n = 0 \qquad (n \in \mathcal{N}(z)).$$

Hence, $\boldsymbol{\xi} = \mathbf{0}$, and Problem (B.8) reduces to

$$\left|\begin{array}{ll} \max\limits_{\boldsymbol{w}_z, \boldsymbol{\phi}, s, b} & -\dfrac{\gamma}{2}\boldsymbol{w}_z^\top \boldsymbol{w}_z - \dfrac{\boldsymbol{\phi}^\top \boldsymbol{\phi}}{4} + s \\ \text{s.t.} & \boldsymbol{w}_z \geq -\boldsymbol{E}_z^\top \boldsymbol{\phi} + 2s\boldsymbol{R}_z^\top \boldsymbol{u}^*(z) + b(\tilde{\boldsymbol{r}}_z - \rho \mathbf{1}), \\ & \boldsymbol{w}_z \geq \mathbf{0}, \quad b \geq 0. \end{array}\right. \tag{B.9}$$

□

Appendix C Proof of Proposition 1

First, the normalization preserves the descent argument used in Proposition 3.2 of [20]. Because the additional term $\frac{1}{2\gamma}\boldsymbol{y}_z^\top \boldsymbol{y}_z$ is quadratic, the regularized objective consists only of quadratic terms.

$$\frac{1}{2\gamma}(\boldsymbol{y}_z^k)^\top \boldsymbol{y}_z^k + (\boldsymbol{v}^k)^\top \boldsymbol{v}^k = \frac{1}{(\hat{\boldsymbol{u}}^k)^\top \hat{\boldsymbol{u}}^k}\left(\frac{1}{2\gamma}(\hat{\boldsymbol{y}}_z^k)^\top \hat{\boldsymbol{y}}_z^k + (\hat{\boldsymbol{v}}^k)^\top \hat{\boldsymbol{v}}^k\right).$$

Therefore, under the same normalization step as in [20], the regularized objective satisfies the same scaling relation, and hence the same descent inequality as in Proposition 3.2 of [20] follows.

Second, the boundedness of the NL iterates follows immediately from the boundedness of $\mathcal{Y}_z$, which has already been established in the existence argument for Problem (2.29). As $\boldsymbol{y}_z^k \in \mathcal{Y}_z$ for all k, the sequence $\{\boldsymbol{y}_z^k\}$ is bounded, and thus so are $\{\boldsymbol{u}^k\}$ and $\{\boldsymbol{v}^k\}$ because $\boldsymbol{u}^k = \boldsymbol{R}_z \boldsymbol{y}_z^k$ and $\boldsymbol{v}^k = \boldsymbol{E}_z \boldsymbol{y}_z^k$.

Finally, when $\varepsilon = 0$, every accumulation point satisfies the KKT conditions of Problem (2.29). The proof follows Proposition 3.3 of [20] with only minor modifications. In particular, from the first-order optimality conditions in a similar manner to Appendix B; see ($B.4a$)–($B.4c$), the stationarity condition with respect to $\boldsymbol{y}_z$ can be written as follows, where the multipliers $\boldsymbol{\pi}$ and $\boldsymbol{\xi}$ are retained to make the stationarity condition explicit, and $\boldsymbol{u}^*(z)$ is replaced by the variable $\boldsymbol{u}$:

$$\frac{1}{\gamma}\boldsymbol{y}_z = -2\boldsymbol{E}_z^\top \boldsymbol{v} + 2s\boldsymbol{R}_z^\top \boldsymbol{u} + b(\tilde{\boldsymbol{r}}_z - \rho \mathbf{1}) + \boldsymbol{\pi} - \boldsymbol{\xi}.$$

Here, the ℓ_2 regularization contributes the additional term $\gamma^{-1}\boldsymbol{y}_z$. Thus, every accumulation point satisfies the KKT conditions of Problem (2.29).

Therefore, the generated sequence is bounded, and every accumulation point satisfies the KKT conditions of Problem (2.29) when $\epsilon = 0$.

□

Appendix D Additional results of algorithm performance

Table 1 gives the optimal objective values for $(N, K, S) = (25, 3, 3)$. The objective values on all seven instances increase with a decrease in γ.

Table 1. Each optimal objective value with $(N, K, S) = (25, 3, 3)$.

γ \ Instance	1	2	3	4	5	6	7
$\frac{0.01}{\sqrt{N}}$	204.69	219.52	226.22	235.51	240.41	262.48	262.41
$\frac{0.1}{\sqrt{N}}$	21.35	22.78	23.47	24.44	24.87	27.11	27.09
$\frac{1}{\sqrt{N}}$	3.01	3.1	3.2	3.33	3.32	3.57	3.53
$\frac{10}{\sqrt{N}}$	1.16	1.13	1.17	1.22	1.15	1.21	1.17
$\frac{100}{\sqrt{N}}$	0.85	0.91	0.91	0.96	0.9	0.9	0.88

Appendix E Investment performance for $N = 200$

Table 2 shows that for $N = 200$, where overfitting is limited, the unregularized case ($\gamma = \infty$) yields the best out-of-sample performance.

Table 2. Investment performance with $(N, K) = (200, 6)$.

S	Metric\\γ	$\frac{0.01}{\sqrt{N}}$	$\frac{0.1}{\sqrt{N}}$	$\frac{1}{\sqrt{N}}$	$\frac{5}{\sqrt{N}}$	$\frac{10}{\sqrt{N}}$	$\frac{100}{\sqrt{N}}$	∞
10	$\mu R^2(\boldsymbol{x})(\%)$	10.1	10.93	14.17	21.74	27.14	39.94	42.22
	Return(%)	25.67	16.85	21.64	21.69	25.37	16.22	24.38
	Risk(%)	28.62	19.96	19.09	20.89	24.98	17.37	21.24
	SR	0.9	0.84	1.13	1.04	1.02	0.93	1.15
	UR(%)	2.9	1.53	1.67	1.77	2.56	1.22	2.28
20	$\mu R^2(\boldsymbol{x})(\%)$	9.71	10.93	13.08	18.49	28.83	39.67	42.25
	Return(%)	22.55	18.66	19.94	21.53	20.49	16.07	24.16
	Risk(%)	24.12	19.18	19.24	19.65	21.38	16.91	21.23
	SR	0.93	0.97	1.04	1.1	0.96	0.95	1.14
	UR(%)	2.26	1.32	1.42	1.59	1.8	1.19	2.26
30	$\mu R^2(\boldsymbol{x})(\%)$	9.63	10.37	12.4	17.98	27.19	39.63	42.25
	Return(%)	21.69	18.5	17.78	20.7	20.36	16.14	24.16
	Risk(%)	23.01	19.14	19.18	18.6	20.9	16.88	21.23
	SR	0.94	0.97	0.93	1.11	0.97	0.96	1.14
	UR(%)	2.08	1.28	1.2	1.42	1.69	1.18	2.26

AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)