CASE STUDY: Support Vector Machines Based on Tail Risk Measures (quadratic, cvar_risk, var_risk, max_cvar_risk, max_var_risk).

Background

This case study illustrates the application of the risk management framework to the Support Vector Machine (SVM) classification problem.

Given a training data \((\xi_1, y_1), (\xi_2, y_2), \ldots, (\xi_J, y_J)\), where \(\xi_j \in \mathbb{R}^n\) are features and \(y_j \in \{-1, 1\}\) are class labels, the basic idea of SVM is to find an optimal separating hyper-plane in the features space which maximizes the margin between two classes. Cortes et al. (1995) proposed to solve SVM classification problem as a quadratic programming. An alternative formulation, known as nu-SVM, was suggested by Scholkopf, et al. (2000). Takeda and Sugiyama (2008) proposed to use the CVaR risk measure in classification and formulated the SVM learning problem as a CVaR minimization problem. Wang (2009) proposed robust nu-Support Vector Machine based on worst-case CVaR Minimization. Tsyurmasto and Uryasev (2012) proposed Support Vector Machines based on Value-at-Risk (VaR) measure. They obtained new SVM classifiers based on VaR risk measure for the following CVaR-based SVMs: Nu-SVM, Extended Nu-SVM, Robust Nu-SVM.

Case study contains the following problem formulations: 1) regularized CVaR, 2) regularized VaR, 3) CVaR minimization with unity constraint, 4) VaR minimization with unity constraint, 5) regularized robust CVaR minimization, 6) regularized robust VaR minimization. Problems 1,2,5,6 include additional quadratic regularization term.

References


Notations

- \(J = \) number of observations (interpreted as scenarios having equal probabilities);
- \(j = \) index of observation points \(\{1, \ldots, J\}\);
- \((\xi_1, y_1), (\xi_2, y_2), \ldots, (\xi_J, y_J)\) = training data, where \(\xi_j \in \mathbb{R}^n\) is the feature column vector and \(y_j \in \{-1, 1\}\) is class label for observation \(j\);
- \(\xi_j^T = \) transposed vector \(\xi_j\);
- \(x = (x', x_0) = \) vector of decision variables;
- \(L(x) = \) random loss function given by observations (scenarios)
  \(L(x) = -y_j(\xi_j^T x' + x_0), j = 1, \ldots, J;\)
- \(M = \) number of data subsets;
- \(J_m = \) number of observation in the data subset \(m\);
- \(cvar_{\alpha}(L(x)) = \) CVaR risk with confidence level \(\alpha\) for random loss function \(L(x)\);
- \(var_{\alpha}(L(x)) = \) VaR risk with confidence level \(\alpha\) for random loss function \(L(x)\);
\text{max}_\text{cvar\_risk}_a(L(x)) = \text{maximum of CVaRs with confidence level } \alpha \text{ over set of data subsets, } m = 1, \ldots, M.

\text{max}_\text{var\_risk}_a(L(x)) = \text{maximum of VaRs with confidence level } \alpha \text{ over set of data subsets, } m = 1, \ldots, M.

\textbf{Optimization Problem 1a}

\textit{Minimize sum of regularization term and CVaR}

\begin{equation}
\min_x \quad (x')^T x' + \text{cvar\_risk}_a(L(x)) \tag{CS.1a}
\end{equation}

\textbf{Optimization Problem 1b}

\textit{Minimize sum of regularization term and VaR}

\begin{equation}
\min_x \quad (x')^T x' + \text{var\_risk}_a(L(x)) \tag{CS.1b}
\end{equation}

\textbf{Optimization Problem 2a}

\textit{Minimize CVaR, subject to unity constraint on quadratic term}

\begin{equation}
\min_x \quad \text{cvar\_risk}_a(L(x)) \quad \text{s.t.} \quad (x')^T x' = 1 \tag{CS.2a}
\end{equation}

\textbf{Optimization Problem 2b}

\textit{Minimize VaR, subject to unity constraint on quadratic term}

\begin{equation}
\min_x \quad \text{var\_risk}_a(L(x)) \quad \text{s.t.} \quad (x')^T x' = 1 \tag{CS.2b}
\end{equation}

\textbf{Optimization Problem 3a}

\textit{Minimize sum of regularization term and maximum of CVaRs}

\begin{equation}
\min_x \quad (x')^T x' + \text{max\_cvar\_risk}_a(L(x)) \tag{CS.3a}
\end{equation}

\textbf{Optimization Problem 3b}

\textit{Minimize sum of regularization term and maximum of VaRs}

\begin{equation}
\min_x \quad (x')^T x' + \text{max\_var\_risk}_a(L(x)) \tag{CS.3b}
\end{equation}