Introduction to large MIMO

Introduction to support vector regression

Firmly grounded in framework of statistical learning theory, Support Vector Machine (SVM) is proposed in 1960s [ref vapnik], and of immense research and industry interest since 1990s. SVM is a powerful tool for supervised learning tasks such as classification, regression and prediction. Moreover, the kernel trick [ref learning with kernels SVM regularization] makes it possible to map data samples into higher dimensional feature space. Therefore SVM can deal with non-linear learning tasks. This makes SVM become a promising tool for complex real-world problems.

Based on the similar principle, epsilon-Support Vector Regression (epsilon-SVR) [vapnik 1995, smola 2003], is developed.

Like SVM, epsilon-SVR first change primal objective function into dual optimization task, then solving the dual quadratic optimization problem. Typically this kind of problem can be solved by numerical quadratic optimization (QP) methods, however, they are computational costly. Decomposition methods, denotes a set of algorithms that divide the optimization problem into sub problems. In each iteration, only a subset of Lagrange multipliers (also named work set) is optimized. Sequential Minimal Optimization (SMO) [ref A fast algorithm sequential minimal optimization] is an extreme case of decomposition methods which chooses dual Lagrange multiplier to optimize in each iteration, SMO uses analytical QP step, makes the solver works much more faster than numerical QP algorithms.

Decomposition methods can be employed to epsilon SVR by the similar manner.

Bouloulis et employ Wirtinger’s calculus into Reproducing Kernel Hilbert Space (RKHS) so that expands real-SVM to pure complex SVM by exploiting complex kernel [ref complex support vector machine]. Based on this work, we construct a prototype of a complexity & performance controllable detector for large MIMO based on dual channel complex SVR. The detector can be divided into two parallel real SVR optimization problem which can be solved independently. Moreover, only real part of kernel matrix is needed in both channel. This means a large amount of computation can be reduced.

Steinwart et[ref SVM without offset] shows with a proper designed work set selection strategy, the approach that choosing double Lagrange multipliers can be much more faster than choosing single Lagrange multiplier without performance loss.

Based on the discrete time MIMO channel model, In our regression model, this CSVR-detector is constructed without offset, The offset in SVR imposes an additional linear quality constraint, which makes it necessary for decomposition methods such as Sequential Minimal Optimization to update more than one Lagrange multipliers in each iteration.

Therefore, for each real SVR without offset, in principle, only one offset is needed to be updated in each iteration, In our prototype, we propose a sequential single Lagrange multiplier search strategy that find two Lagrange multiplier sequentially, which can approximate the optimal dual Lagrange multiplier searching strategy. The former one only requires O(n) searches in one iteration, while the optimal dual Lagrange multiplier strategy requires O(n2) searches per iteration.