Abstract: Many modern big data applications feature large scale in both numbers of responses and predictors. Better statistical efficiency and scientific insights can be enabled by understanding the large-scale response-predictor association network structures via layers of sparse latent factors ranked by importance. Yet sparsity and orthogonality have been two largely incompatible goals. To accommodate both features, in this presentation we first suggest the method of sparse orthogonal factor regression (SOFAR) via the sparse singular value decomposition with orthogonality constrained optimization to learn the underlying association networks, with broad applications to both unsupervised and supervised learning tasks such as biclustering with sparse singular value decomposition, sparse principal component analysis, sparse factor analysis, and spare vector autoregression analysis. Exploiting the framework of convexity-assisted nonconvex optimization, we derive nonasymptotic error bounds for the suggested procedure characterizing the theoretical advantages. The statistical guarantees are powered by an efficient SOFAR algorithm with convergence property. Next, we propose a consistent estimation method for the approximate factor model of Chamberlain and Rothschild (1983, Econometrica), known recently as weak factor models, by extending the SOFAR methodology. In our experiment, the performance of the new estimator dominates that of the principal component estimators in terms of mean absolute loss, and its superiority gets larger as the factor components become weaker. We apply our method to analyse S&P500 firm security monthly returns, and the results show that the largest eigenvalue grows proportional to the number of the firms, whilst the second and third much less slowly diverge.