The key limitation of our spatio-temporal compressive sensing framework is that it only considers missing values and the low-rank structure and does not explicitly account for anomalies or measurement noise. As a result, the performance may degrade in the presence of significant noise or anomalies.

We propose *LENS decomposition*, a general framework for analyzing network matrices by decomposing the matrix into a Low-rank matrix, an Error term, a Noise matrix, and a Sparse matrix. Our formulation is much more general than some of the latest development in compressive sensing (*e.g.*, robust PCA [1]) and has many potential applications. For example, the sparse component is useful for network anomaly detection; the low-rank component is useful for network tomography, interpolation, prediction, and synthesis.

1 LENS Decomposition

Basic formulation. We propose to decompose an $m \times n$ data matrix D into a low-rank matrix X, a sparse matrix Y, a noise matrix Z, and an error matrix W by solving the following *convex* optimization problem:

$$\begin{aligned} \text{minimize:} \quad & \alpha\|X\|_* + \beta\|Y\|_1 + \frac{1}{2\sigma}\|Z\|_F^2, \\ \text{subject to:} \quad & X + Y + Z + E.*W = D, \end{aligned} \tag{1}$$

where:

- X is the low-rank component; $\|X\|_*$ is the nuclear norm of matrix X, which penalizes against high rank of X and can be computed as the total sum of X's singular values.
- Y is the sparse (*i.e.*, anomalous) component; $\|Y\|_1 = \sum_{i,j} |Y[i,j]|$ is the ℓ_1-norm of Y, which penalizes against lack of sparsity in Y.
- Z is a dense noise component; $\|Z\|_F^2$ is the squared Frobenius norm of matrix Z, which penalizes against large entries of Z and can be computed as $\|Z\|_F^2 = \sum_{i,j} Z[i,j]^2$.
- E is a binary error indicator matrix such that $E[i,j] = 1$ iff entry $D[i,j]$ is erroneous or missing. Let $\eta(D) = 1 - \frac{\sum_{i,j} E[i,j]}{m \times n}$ be the fraction of D's elements that are neither missing nor erroneous.
- W is the arbitrary error component, with $W[i,j] \neq 0$ only when $E[i,j] = 1$ (thus $E.*W = W$, where $.*$ is element-wise multiplication). Since W fully captures the erroneous or missing values, we can set $D[i,j] = 0$ whenever $E[i,j] = 1$ without loss of generality.
- σ is the standard deviation of $Z[E = 0] \triangleq \{Z[i,j] \mid E[i,j] = 0\}$. For simplicity, we assume that σ is known *a priori* and that Z is homoscedastic (*i.e.*, with uniform variance).

Generalization. We can easily generalize Eq. (1) to cope with more general measurement constraints:

$$AX + BY + CZ + E.*W = D, \tag{2}$$

where:

- A captures tomographic constraints that involve both direct and indirect measurements.
- B represents an overcomplete anomaly profile matrix. For example, to enumerate all possible spike locations, we can simply set B to be the identity matrix I. Alternatively, B can also be constructed using Haar wavelet transform matrix, or the discrete cosine transform matrix, or a combination of these. We ensure that columns of B are distinct and have unit length.
- C captures the correlation among measurement noise.

In this case, notice that (i) when X is low-rank, AX is also low-rank, and (ii) when Z is a dense noise matrix, CZ is also likely to be dense. We therefore propose to infer X, Y, Z, and W by solving

$$\begin{aligned} \text{minimize:} \quad & \alpha\|AX\|_* + \beta\|Y\|_1 + \frac{1}{2\sigma}\|CZ\|_F^2, \\ \text{subject to:} \quad & AX + BY + CZ + E.*W = D, \end{aligned} \tag{3}$$

where σ becomes the standard deviation of (non-erroneous) elements of CZ instead of Z.

Note that we can simplify Eq. (3) by performing a change of variable. Specifically, let $X = AX_{\text{orig}}$ and $Z = CZ_{\text{orig}}$, then Eq. (3) becomes:

$$\begin{aligned} \text{minimize:} \quad & \alpha\|X\|_* + \beta\|Y\|_1 + \frac{1}{2\sigma}\|Z\|_F^2, \\ \text{subject to:} \quad & X + BY + Z + E. * W = D. \end{aligned} \tag{4}$$

Once we solve Eq. (4), we can then infer X_{orig}, and Z_{orig} according to $X_{\text{orig}} = \mathsf{pinv}(A)X$ and $Z_{\text{orig}} = \mathsf{pinv}(C)Z$, where $\mathsf{pinv}(M)$ gives the pseudoinverse of matrix M.

References

[1] E. J. Candes, X. Li, Y. Ma, and J. Wright. Robust principal component analysis?, 2009. Manuscript. Available from `http://www-stat.stanford.edu/~candes/papers/RobustPCA.pdf`.