

# MODELO DE MEDICIÓN PARA LA ESTIMACIÓN DE LA MATRIZ DE CORRELACIÓN DE DATOS ORDENADOS



**JUAN PABLO HOYOS SANCHEZ**

Universidad del Cauca  
Facultad de Ingeniería Electrónica y Telecomunicaciones  
Departamento de Electrónica y Telecomunicaciones  
Popayan, Colombia  
2018

## Resumen

Los métodos basados en estadísticos de orden o *order statistics*, son a menudo usados en finanzas, calidad de control, procesamiento de datos y señales, especialmente cuando las señales de interés están inmersas en ruido impulsivo. Estos permiten incluir información de tiempo a costa de incrementar la dimensión del problema. En problemas de grandes dimensiones, usualmente requerimos conocer solo estadísticos de segundo orden.

En esta tesis se diseñó un modelo de medición basado en mediciones cuadráticas de rango uno o *sketch* para estimar la matriz de correlación de datos que han sido ordenados. Se estudia la estructura de la matriz para un proceso aleatorio vectorial estacionario, y mostramos que su estructura depende de una submatriz diagonal y una submatriz con diagonal cero. Explotando esta estructura se diseñan dos algoritmos, el primero luego de la estimación realiza un promedio y el segundo algoritmo tiene en cuenta la información de la estructura para la definición de una nueva relajación convexa para recuperar la matriz. Se demuestra que el número de mediciones es proporcional al tamaño original del problema (sin ordenamiento), además se desarrolla un muestreo en bloques que optimiza el muestreo de la estructura de la matriz. Para ambos algoritmos se presentan simulaciones para mostrar el desempeño en la reconstrucción y la robustez de la estimación cuando el ruido uniforme está presente, superando el desempeño mostrado por el estimador referente.

Para los casos en que no es posible suponer un número infinito de realizaciones del proceso estacionario, se desarrolla una modificación al esquema de medición que permite una convergencia más rápida, con lo cual se consolida un modelo de medición para la estimación de la matriz de correlación.

Se desarrolló una extensión del segundo algoritmo basado en métodos de división proximal, en especial en los métodos *forward-backward* y primal-dual, que al solucionar iterativamente una secuencia de subproblemas logra una computación eficiente para el caso de datos de grandes dimensiones.

Por último, basados en el modelo de medición se diseñó un estimador para datos ordenados que presentan un comportamiento no-estacionario, mostrando una importante disminución en el consumo de recursos y una leve reducción en el número de mediciones respecto al estimador referente.

**Palabras clave:** Matriz de correlación, estadísticos de orden, mediciones cuadráticas, *sketch*, métodos de división proximal, *forward-backward*, primal-dual, ruido uniforme.

# Abstract

Methods based on ordered statistics are often used in finance, quality control, data and signals processing, especially when the signals of interest are immersed in impulsive noise. These allow to include information of time at the expense of increasing the dimension of the problem. In large problems, we usually need to know only second-order statistics.

In this thesis a measurement model based on quadratic measurements of rank one or textit sketch was designed to estimate the matrix of correlation of data that have been ordered. We study the structure of the matrix for a stationary vector random process, and show that its structure depends on a diagonal submatrix and a submatrix with zero diagonal. Exploiting this structure two algorithms are designed, the first after the estimation performs an average and the second algorithm takes into account the information of the structure for the definition of a new convex relaxation to recover the matrix. It is demonstrate that the number of measurements is proportional to the original size of the problem (without ordering), in addition a block sampling is developed to optimize the sampling of the structure of the matrix. For both algorithms we provide simulations to show the performance in the reconstruction and the robustness of the estimate when uniform noise is present, surpassing the performance shown by the referent estimator.

For cases in which it is not possible to assume an infinite number of realizations of the stationary process, a modification to the measurement scheme is developed that allows a faster convergence, which consolidates a measurement model for the estimation of the correlation matrix.

We developed an extension of the second algorithm based on proximal division methods, especially in the forward-backward and primal-dual methods, which when iteratively solving a sequence of sub-problems ensures an efficient computation for the case of large data.

Finally, based on the measurement model, an estimator was designed for ordered data that exhibit a non-stationary behavior, showing a significant decrease in the consumption of resources and a slight reduction in the number of measurements with respect to the referent estimator.

**Keywords:** Correlation matrix, ordered statistics, quadratic measurements, sketch, proximal splitting methods, forward-backward, primal-dual, uniform noise.