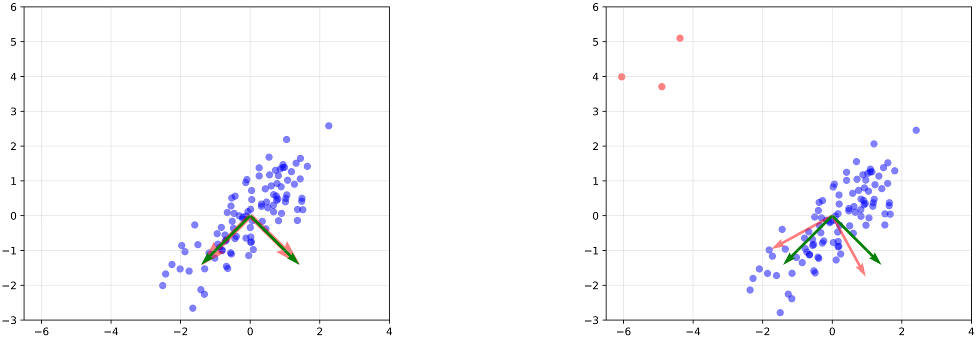
Tyler's M-estimator Through the Lens Of Convex-Concave Programming

Daniel Cederberg

Stanford University

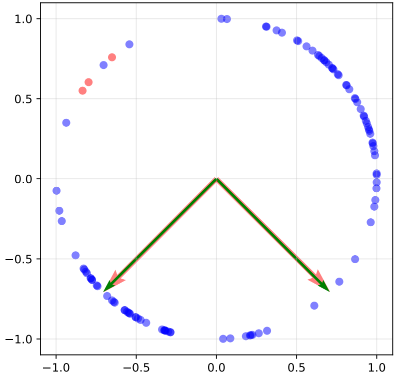
May 2, 2026

Robust covariance estimation



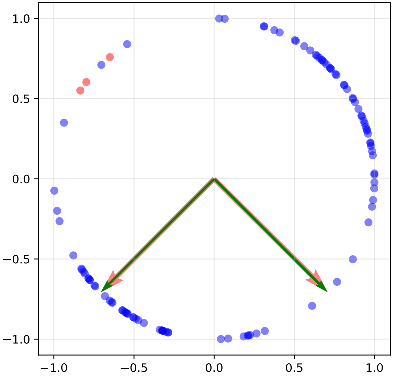
- classical covariance estimators are sensitive to outliers (red points)
- green arrows = true principal components, red = sample principal components

Angular central Gaussian distribution



- normalize samples to the unit sphere: $X \triangleq Z/\|Z\|_2$ where $Z \sim N(0, \Sigma)$
- outliers can only influence direction, not magnitude

Angular central Gaussian distribution



- normalize samples to the unit sphere: $X \triangleq Z/\|Z\|_2$ where $Z \sim N(0, \Sigma)$
- outliers can only influence direction, not magnitude
- density of X is *angular central Gaussian* (ACG):

$$p(x; \Sigma) = \frac{\Gamma(n/2)}{2\pi^{n/2}\sqrt{\det \Sigma}} (x^T \Sigma^{-1} x)^{-n/2}$$

Tyler's M-estimator (Tyler, 1987)

A distribution-free M-estimator of multivariate scatter
DE Tyler - The annals of Statistics, 1987 - JSTOR
The existence and uniqueness of a limiting form of a Huber-type M-estimator of multivariate scatter is established under certain conditions on the observed sample. These conditions …

$$\text{minimize} \log\det\Sigma + \frac{n}{m}\sum_{i=1}^{m}\log(x_i^T\Sigma^{-1}x_i)$$

- variable $\Sigma \in \mathbf{S}^n$, data $x_1, \ldots, x_m \in \mathbf{R}^n$
- MLE for the angular central Gaussian distribution
- well-known in the statistics literature (industry adoption is another story?)
- computed via fixed-point iteration:

$$\Sigma_{k+1} = \frac{n}{m}\sum_{i=1}^{m}\frac{x_i x_i^T}{x_i^T\Sigma_k^{-1}x_i}$$

Does Tyler's fixed-point iteration converge?

- non-convex problem, but global *asymptotic* convergence known (Kent & Tyler, 1988)

Does Tyler's fixed-point iteration converge?

- non-convex problem, but global *asymptotic* convergence known (Kent & Tyler, 1988)
- we establish the first global *non-asymptotic* rate: $\mathcal{O}(1/k)$
  - key idea: reinterpret fixed-point iteration as the convex-concave procedure
  - enables an optimization (rather than fixed-point) analysis
  - analysis combines Riemannian optimization with relative smoothness

Does Tyler's fixed-point iteration converge?

- non-convex problem, but global *asymptotic* convergence known (Kent & Tyler, 1988)
- we establish the first global *non-asymptotic* rate: $\mathcal{O}(1/k)$
  - key idea: reinterpret fixed-point iteration as the convex-concave procedure
  - enables an optimization (rather than fixed-point) analysis
  - analysis combines Riemannian optimization with relative smoothness
- come by my poster if you like covariance matrices, convergence rates (or CVXPY)