Contribution associée

MAXIMUM LIKELIHOOD ESTIMATION IN PRINCIPAL COMPONENTS ANALYSIS WITH NOISE PERTURBATION : A GAUSSIAN MODEL

Jean-Louis PHILOCHE (*), Damien WYART (**)

(*) Groupe des écoles des télécommunications,
(**) Université d'Evry

1 Theorem

We give a new proof of a Theorem which explicit the (M.L.E.) out of an iid sample of size n :
$(y_1, \ldots, y_n)$, each y_i follows $\mathcal{N}_p(m,\Sigma)$.

$$\textit{Assumption} : \qquad l_1 > l_2 > \ldots > l_q > l_{q+1} = \ldots = l_p > 0 \qquad (1)$$

The (l_j) are the (unknown) eigenvalues of Σ. For the (y_i) given, the log-likelihood is a function of m and Σ. The log-likelihood depends of the data only through $\bar{y}$ and S (the sample covariance). We can limit ourselves to the case $m = 0$, so : $L(\Sigma) \coloneqq -\ln(\det(\Sigma)) - \mathrm{Tr}(S\Sigma^{-1})$, $(\Sigma \in \mathcal{U})$. $\mathcal{U}$, subset of $\mathcal{M}^+$ (the set of the symmetric and positive definite operators (in $\mathbb{R}^p$) is characterised, in $\mathcal{M}^+$, by the fact that every Σ in $\mathcal{U}$ satisfies (1).

Theorem (I) : On $\mathcal{U}$, L attains a maximum at an unique point $\hat{\Sigma}$, where the corresponding eigenvalues $(\hat{l}_j)$ are $\hat{l}_j = s_j (1 \leq j \leq q)$, and $\hat{l}_j = \frac{1}{p-q} \sum_{k=q+1}^{p} s_k$, $(1 \leq j \leq q)$. (Clearly, the eigenvalues of S are distinct and strictly positive (a.s), so we can suppose : $s_1 > s_2 > \ldots > 0$).

(II) Moreover, any maximising sequence (Σ_n) in $\mathcal{U}$, i.e. such that $\lim_{n\to\infty} L(\Sigma_n)=L(\hat{\Sigma})$, converges to $\hat{\Sigma}$.

2 Principal means to review to build a proof

(i) We shall use a *change of variables* : $\Theta = \Sigma^{-1}$, $\mathcal{N} = \{\Theta \in \mathcal{M}^+ : \Theta^{-1} \in \mathcal{U}\}$, with $F(\Theta) = L(\Theta^{-1}) \coloneqq \ln(\det(\Theta)) - \mathrm{Tr}(S\Theta)$. It is very convenient to work with F which is *strictly concave*.

(ii) We also introduce the action of the group $\mathcal{SO}(p)$ on $\mathcal{GL}(p)$: this conjugation is defined by : $(U,\Sigma) \to U\Sigma U^{-1}$, with $U \in \mathcal{SO}(p)$, $S \in \mathcal{GL}(p)$.

As $\mathcal{N}$ is an union of orbits of the action of $\mathcal{SO}(p)$, to find the maximum of F in $\mathcal{N}$, we can, first, find it on any orbit, then find the maximum of the maxima.