

TAKE-HOME MESSAGE

In the general context of linear inverse problems with colored Gaussian noise, our results prove that **the Bayesian conditional mean estimator is always the solution of a penalized regression problem.**

“PRIORS” AND “PRIORS”

$$y = \mathbf{A}z + b, \quad \text{with } z \in \mathbb{R}^D, y \in \mathbb{R}^n, \mathbf{A} \in \mathbb{R}^{n \times D}, P_B \sim \mathcal{N}(0, \Sigma).$$

Bayesian estimation:

$$\psi_{\mathbf{A}, \Sigma, P_Z}(y) := \mathbb{E}(Z|Y = y)$$

n.b.: explicitly relies on P_Z (“Bayesian prior”)

Regularization-based estimation:

$$\psi(y) := \operatorname{argmin}_{z \in \mathbb{R}^D} \frac{1}{2} \|y - \mathbf{A}z\|^2 + \phi(z)$$

n.b.: explicitly relies on ϕ (“regularization prior”)

⇒ Do they provide different estimators? How do these “priors” relate to each other?

MAP INTERPRETATION

$$\text{Maximum A Posteriori with prior } e^{-\phi(z)} \Leftrightarrow \operatorname{argmin}_z \frac{1}{2} \|y - \mathbf{A}z\|^2 + \phi(z) \quad (\text{PLS})$$

Is (PLS) specifically cut for unknowns with distribution $e^{-\phi}$?

SPECIAL CASE [1]: WHITE GAUSSIAN DENOISING

From [1], for $\mathbf{A} = \mathbf{I}$ and $\Sigma = \mathbf{I}$, i.e.: $y = z + b$

Theorem (Main result of [1]). *For any non-degenerate prior P_Z , we have:*

1. $\psi_{\mathbf{I}, \mathbf{I}, P_Z}$ is one-to-one;
2. $\psi_{\mathbf{I}, \mathbf{I}, P_Z}$ and its inverse are C^∞ ;
3. $\forall y \in \mathbb{R}^n, \psi_{\mathbf{I}, \mathbf{I}, P_Z}(y)$ is the unique global minimum and unique stationary point of

$$z \mapsto \frac{1}{2} \|y - z\|^2 + \phi_{\mathbf{I}, \mathbf{I}, P_Z}(z), \quad \text{with:} \quad (1)$$

$$\phi_{\mathbf{I}, \mathbf{I}, P_Z}(z) := \begin{cases} -\frac{1}{2} \|\psi_{\mathbf{I}, \mathbf{I}, P_Z}^{-1}(z) - z\|_2^2 - \log p_Y[\psi_{\mathbf{I}, \mathbf{I}, P_Z}^{-1}(z)]; & \text{for } z \in \operatorname{Im} \psi_{\mathbf{I}, \mathbf{I}, P_Z}; \\ +\infty, & \text{for } z \notin \operatorname{Im} \psi_{\mathbf{I}, \mathbf{I}, P_Z}; \end{cases}$$

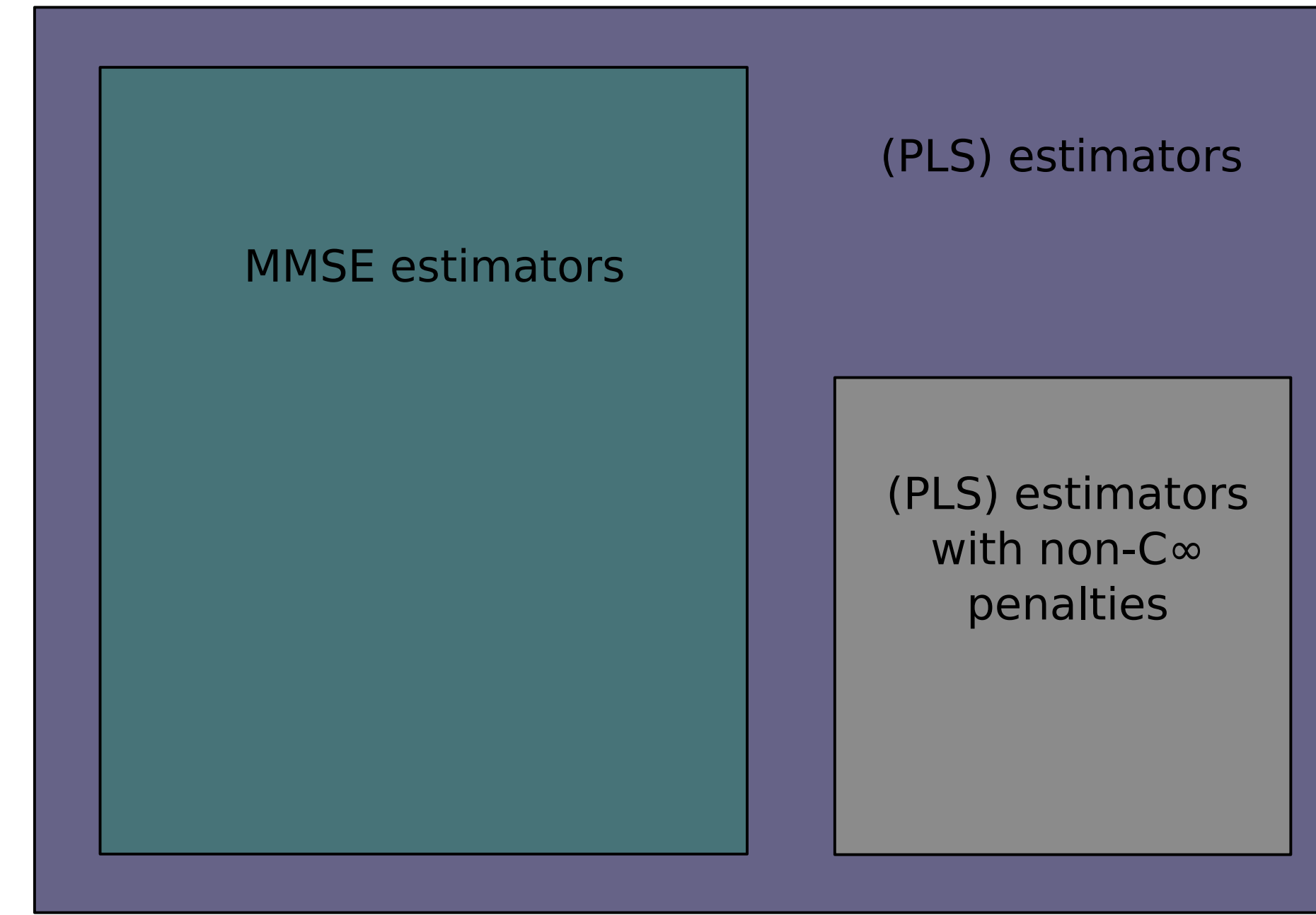
4. The penalty function $\phi_{\mathbf{I}, \mathbf{I}, P_Z}$ is C^∞ ;
5. Any penalty function $\phi(z)$ such that $\psi_{\mathbf{I}, \mathbf{I}, P_Z}(y)$ is a stationary point of (1) satisfies $\phi(z) = \phi_{\mathbf{I}, \mathbf{I}, P_Z}(z) + C$ for some constant C and all z .

THE MAIN RESULT

Theorem (Main result). *Let $h : \mathbb{R}^{(D-n)} \times \mathbb{R}^{(D-n)} \mapsto \mathbb{R}^+$ be any function such that $h(x_1, x_2) = 0 \Rightarrow x_1 = x_2$. For any non-degenerate prior P_Z , any non-degenerate Σ and \mathbf{A} , we have:*

1. $\psi_{\mathbf{A}, \Sigma, P_Z}$ is injective.
2. $\forall y \in \mathbb{R}^n, \psi_{\mathbf{A}, \Sigma, P_Z}$ is the unique global minimum and stationary point of $z \mapsto \frac{1}{2} \|y - \mathbf{A}z\|_\Sigma^2 + \phi_{\mathbf{A}, \Sigma, P_Z}^h(z)$.
3. $\phi_{\mathbf{A}, \Sigma, P_Z}^h$ is C^∞ if and only if h is.

CONSEQUENCES



RELATIONS TO THE EVIDENCE

Lemma 1 (Convexity and Separability). *For any non-degenerate prior P_Z , Theorem says that $\forall y \in \mathbb{R}^n, \psi_{\mathbf{I}, \mathbf{I}, P_Z}(y)$ is the unique global minimum and stationary point of $z \mapsto \frac{1}{2} \|y - \mathbf{I}z\|^2 + \phi_{\mathbf{I}, \mathbf{I}, P_Z}(z)$. Moreover, the following results hold:*

1. $\phi_{\mathbf{I}, \mathbf{I}, P_Z}$ is **convex** if and only if $p_Y(y) := p_B \star P_Z(y)$ is **log-concave**,
2. $\phi_{\mathbf{I}, \mathbf{I}, P_Z}$ is **additively separable** if and only if $p_Y(y)$ is **multiplicatively separable**.

FUTURE WORKS

- Generalizing to other classes of problems
- For under-determined problems, can we leverage the additional degrees of freedom?
- Non-Gaussian noise? Different choice of loss?
- Can we come up with the metric in the estimators space?

(VERY) SELECTED BIBLIOGRAPHY

- [1] Rémi Gribonval. Should penalized least squares regression be interpreted as maximum a posteriori estimation? *IEEE Transactions on Signal Processing*, 59(5):2405–2410, 2011.

ACKNOWLEDGEMENTS

This work is supported in part by the European Research Council, PLEASE project (ERC-StG-2011-277906)

BERNOULLI-GAUSSIAN PRIOR P_Z

$$p_Z(z) = p \frac{1}{2\pi\sigma_0} e^{-\frac{z^2}{2\sigma_0^2}} + (1-p) \frac{1}{2\pi\sigma_1} e^{-\frac{z^2}{2\sigma_1^2}}, \quad \text{with } \sigma_0^2 \rightarrow 0$$

