Internship proposal
Sparse regression and dimension reduction
Application to sensor measurements and data normalization

IFP Energies nouvelles, Rueil-Malmaison (Paris suburbs), France

Subject

The instrumental context is that of multiple 1D data or measurements y_m related to the the same phenomenon x, corrupted by random effects n_m and a different scaling parameter a_m, due to uncontrolled sensor calibrations or measurement variability. The model is thus:

$$y_m(k) = a_m x(k) + n_m(k) \,.$$

The aim of the internship is to robustly estimate scaling parameters a_m (with confidence bounds) in the presence of missing data or outliers for potentially small, real-life signals x with large amplitude variations. The estimation should be as automatized as possible, based on data properties and priors (*e.g.* sparsity, positivity), so as to be used by non-expert users. Signals under study are for instance: vibration, analytical chemistry or biological data.

Of particular interest for this internship is the study and performance assessment of robust loss or penalty functions (around the $\ell_{2,1}$-norm) such as the R_1-PCA [DZHZ06] or [NHCD10] and low-rank decomposition as in GoDec [ZT11].

Education profile

Second/third year engineering school and/or master of science with strong skills signal/image processing, optimization, statistics, machine learning, applied mathematics. Applicants should provide a resume and a motivation letter emphasizing prior knowledge related to the subject.

Information updates

`http://www.laurent-duval.eu/intern-sparse-regression-dimension-reduction.html`

Internship tutor

Laurent Duval, (Control, signal and systems, IFP Energies nouvelles); laurent.duval[]ifpen.fr

References

[DZHZ06] Chris Ding, Ding Zhou, Xiaofeng He, and Hongyuan Zha. R_1-PCA: Rotational invariant L_1-norm principal component analysis for robust subspace factorization. In *Proc. Int. Conf. Mach. Learn.*, Pittsburgh, PA, USA, Jun. 25-29, 2006. Association for Computing Machinery (ACM).

[NHCD10] Feiping Nie, Heng Huang, Xiao Cai, and Chris H. Ding. Efficient and robust feature selection via joint $\ell_{2,1}$-norms minimization. In *Proc. Ann. Conf. Neur. Inform. Proc. Syst.*, 2010.

[ZT11] Tianyi Zhou and Dacheng Tao. GoDec: Randomized low-rank & sparse matrix decomposition in noisy case. In *Proc. Int. Conf. Mach. Learn.*, 2011.