Non-negative Matrix Factorization with Gaussian Process Priors

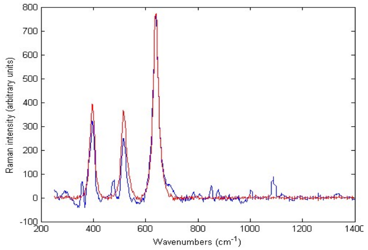
Technical University of Denmark

DTU Compute
Institut for Matematik og Computer Science

Elika Araghi, Sanaz Behboodi, Christian A. Rasmussen

Raman Spectroscopy

- Gaussian like
- Non-negative



NMF decomposition

$$X = DH + N$$

Gaussian process priors

- Link function
- Covarince Function

Maximum a Posteriori NMF

$$P_{D,H|X}(\mathrm{D,H|X}) = \frac{P_{D,H|X}(\mathrm{D,H|X})\, P_{D,H}(\mathrm{D,H})}{P_X(X)}$$

$$\mathcal{L}_{D,H|X}(\mathrm{D,H}) \propto \mathcal{L}_{X|D,H}(D,H) + \mathcal{L}_{D,H}(D,H)$$

Sampling

- Metropolis Hastings

$$A_k(z^*, z^{(\tau)}) = min\left(1, \frac{\tilde{p}(z^*)q_k(z^{(\tau)}|z^*)}{\tilde{p}(z^{(\tau)})q_k(z^*|z^{(\tau)})}\right)$$

Gaussian Process Priors

Covariance Function and Link Function

- Covariance function as a Gaussian radial basis function (RBF)

$$\phi(i,j) = exp\left(-\frac{(i-j)^2}{\beta^2}\right)$$

- For H and D, the inverse exponential link functions is

$$f_h^{-1} = -\frac{1}{\lambda} log\left(\frac{1}{2} - \frac{1}{2}\Phi\left(\frac{h_i}{\sqrt{2}\sigma_i}\right)\right)$$

Change of Optimization Variable

$$H = g_h(\boldsymbol{\eta}) = \mathrm{vec}^{-1}\left(f_h^{-1}(C_h^T\boldsymbol{\eta})\right)$$

$$p_\eta(\boldsymbol{\eta}) = p_H(g_h(\boldsymbol{\eta}))|\mathcal{J}(g_h(\boldsymbol{\eta}))| = \frac{1}{(2\pi)^{LM/2}} \exp\left(-\frac{1}{2}\boldsymbol{\eta}^T\boldsymbol{\eta}\right)$$

- The negative log prior: $\mathcal{L}_\eta(\boldsymbol{\eta}) \propto \frac{1}{2}\boldsymbol{\eta}^T\boldsymbol{\eta}$
- Compute the MAP estimate of the transformed variables

$$\mathcal{L}_{\delta,\eta|X}(\delta,\eta) = \mathcal{L}_{X|D,H}(g_d(\delta), g_h(\eta)) + \frac{1}{2}\delta^T\delta + \frac{1}{2}\eta^T\eta$$

$$\{\delta_{MAP}, \eta_{MAP}\} = \underset{\delta,\eta}{\mathrm{argmin}}\, \mathcal{L}_{\delta,\eta|X}(\delta,\eta)$$

- Using the least squares likelihood, the posterior distribution is

$$\mathcal{L}^{LS-GPP}_{\delta,\eta|X}(\delta,\eta) = \frac{1}{2}\left(\sigma_N^{-2}\|X - g_d(\delta)g_h(\eta)\|_F^2 + \delta^T\delta + \eta^T\eta\right)$$

Results

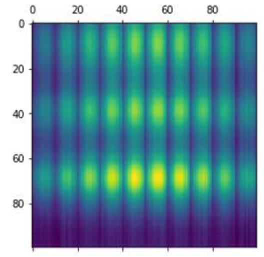


Fig. 1. Generated data simulating a Raman a spectrum

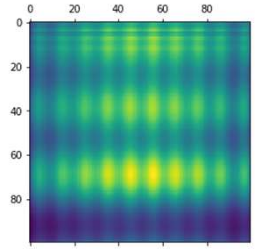


Fig. 2. Reconstructed data of the Raman spectrum

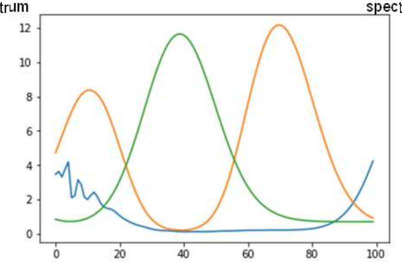


Fig. 3. Features of D. Notice how peaks are smooth and exponential as given by the underlying Gaussian process prior

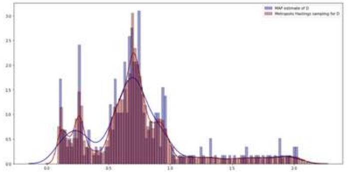


Fig. 5. Histogram comparing the distributions of D (red) and samples (blue)

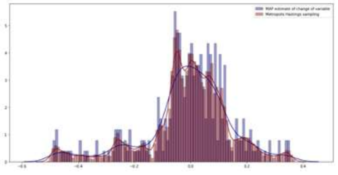


Fig. 4. Histogram comparing the distributions of δMAP (red) and samples (blue)

Conclusion

- Success: Peak- and voigt locations
- Problems: Acceptance criterion by MAP estimate.

References:
[1] Mikkel N. Schmidt and Hans Laurberg, "Non-negative matrix factorization with gaussian process priors," in Computational Intelligence and Neuroscience, 2008.