Correlated models

Billy Fang

October 14, 2016

1 General case

The reconstruction loss term in the loss will stay the same. [Is this correct?]
In general, the KL-divergence term in the loss is

$$D_{KL}(N(\mu, \Sigma) \| N(0, I_k)) = \frac{1}{2}\big(\operatorname{tr}\Sigma + \mu^\top \mu - k - \log\det\Sigma\big).$$

2 Diagonal case

We have already seen in the `jmetzen` tutorial that for $\Sigma = \operatorname{diag}(\sigma_1^2, \ldots, \sigma_k^2)$, we have

$$D_{KL}(N(\mu, \Sigma) \| N(0, I_k)) = \frac{1}{2}\sum_{i=1}^k (\sigma_i^2 + \mu_i^2 - 1 - \log\sigma_i^2)$$

3 Covariances differ across disjoint blocks

If $k = w^2$ and divide a $w \times w$ image into $(w/s)^2 = k/s^2$ disjoint squares J of size $s \times s$ and declare each square J to have covariance structure $\Sigma_J := \operatorname{diag}(\sigma_i^2 - \rho_J)_{i \in J} + \rho_J \mathbf{1}_{s^2}\mathbf{1}_{s^2}^\top$. By the matrix determinant lemma, the determinant of this $s^2 \times s^2$ matrix is

$$\det\Sigma_J = \left(1 + \rho_J \sum_{i \in J} \frac{1}{\sigma_i^2 - \rho_J}\right) \prod_{i \in J}(\sigma_i^2 - \rho_J).$$

Thus, $\Sigma = \operatorname{diag}(\Sigma_{J_1}, \ldots, \Sigma_{J_{k/s^2}})$ is block diagonal with k/s^2 blocks of size $s^2 \times s^2$, and we have

$$\begin{aligned} &D_{KL}(N(\mu, \Sigma) \| N(0, I_k)) \\ &= \frac{1}{2}\left(\sum_{i=1}^k (\sigma_i^2 + \mu_i^2 - 1) - \sum_J \left[\log\left(1 + \rho_J \sum_{i \in J}\frac{1}{\sigma_i^2 - \rho_J}\right) + \sum_{i \in J}\log(\sigma_i^2 - \rho_J)\right]\right) \end{aligned} \tag{1}$$

In the special case where $\rho_J = \rho$ is the same for all squares, the above can be rewritten as

$$D_{KL}(N(\mu, \Sigma) \| N(0, I_k)) = \frac{1}{2}\left(\sum_{i=1}^k (\sigma_i^2 + \mu_i^2 - 1 - \log(\sigma_i^2 - \rho)) - \sum_J \log\left(1 + \rho\sum_{i \in J}\frac{1}{\sigma_i^2 - \rho}\right)\right).$$

Implementation: your $w \times w$ image is represented as a vector of length $k = w^2$ by concatenating the rows of the image. You have a vector of variances $\sigma^2 = (\sigma_i^2)_{i=1}^k$ and block correlations $\rho = (\rho_J)_{J=1}^{k/s^2}$.

Let A be the $k \times (k/s^2)$ matrix whose Jth column is the indicator vector for which of the $k = w^2$ pixels correspond to the square J. For example, if $k = w^2 = 4^2$ and $s = 2$, the transpose is

$$A^\top = \begin{bmatrix} 1 & 1 & & & & & & & 1 & 1 & & & & & & \\ & & 1 & 1 & & & & & & & 1 & 1 & & & & \\ & & & & 1 & 1 & & & & & & & 1 & 1 & & \\ & & & & & & 1 & 1 & & & & & & & 1 & 1 \end{bmatrix}$$

The matrix A can be constructed using the following.

```
A = np.kron(np.resize(np.eye((w/s)**2),(w**2/s,(w/s)**2)), np.ones((s,1)))
```

Then we may form a k-dimensional vector ν with entries $\sigma_i^2 - \rho_J$ by taking $\sigma^2 - A\rho$. Another vector γ with entries $\frac{\rho_J}{\sigma_i^2 - \rho_J}$ can be obtained by taking $A\rho \odot (1/\nu)$.

```
nu = sigma_sq - A.dot(rho)
gamma = A.dot(rho) * (1/nu)
```

Then $A^\top \gamma$ is a vector of length k/s^2 with entries $\rho_J \sum_{i \in J} \frac{1}{\sigma_i^2 - \rho}$.

Then, the KL divergence (1) is as follows.

```
latent_loss = 0.5 * np.sum(sigma_sq + np.square(mu) - 1)
latent_loss -= 0.5 * np.sum(np.log(1+A.T.dot(gamma)))
latent_loss -= 0.5 * np.sum(nu)
```

4 Common local covariance structure

4.1 Horizontal neighbors

Suppose each pixel has covariance ρ with each of the pixels immediately on its left and right. Then the covariance matrix Σ is tridiagonal. More specifically, it is block diagonal with w tridiagonal blocks of of size w^2. Let these blocks be denoted $\Sigma_1, \ldots, \Sigma_w$. For example with $w = 3$, the first block is

$$\Sigma_1 = \begin{bmatrix} \sigma_1^2 & \rho & \\ \rho & \sigma_2^2 & \rho \\ & \rho & \sigma_3^2 \end{bmatrix}.$$

Then

$$D_{KL}(N(\mu, \Sigma) \| N(0, I_k)) = \frac{1}{2}\left(\sum_{i=1}^{k} (\sigma_i^2 + \mu_i^2 - 1 - \log(\sigma_i^2 - \rho)) - \sum_{j=1}^{w} \log \det \Sigma_j \right).$$

One can compute each determinant in $O(w)$ time.

4.2 Vertical neighbors

Suppose each pixel has covariance ρ with each of the pixels immediately above and below it. In the case $w = 3$, we have

$$\Sigma = \begin{bmatrix} \mathrm{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2) & \rho I_3 & 0 \\ \rho I_3 & \mathrm{diag}(\sigma_4^2, \sigma_5^2, \sigma_6^2) & \rho I_3 \\ 0 & \rho I_3 & \mathrm{diag}(\sigma_7^2, \sigma_8^2, \sigma_9^2) \end{bmatrix}$$

In general, Σ is "block tridiagonal" where the on-diagonal blocks are diagonal and simply contain the variance entries, and where the off-diagonal blocks are ρI_w. However, after permuting the components, this can be made into a tridiagonal matrix as well, since this model is not any different than the previous one.

4.3 Four immediate neighbors

Now suppose we combine the above two models: each pixel is correlated with its four immediate neighbors. Then with Σ_i defined as above, we have

$$\Sigma = \begin{bmatrix} \Sigma_1 & \rho I_3 & 0 \\ \rho I_3 & \Sigma_2 & \rho I_3 \\ 0 & \rho I_3 & \Sigma_3 \end{bmatrix}$$

It is clear that this is not a pentadiagonal matrix, since two neighboring pixels do not have common neighbors.