T-61.5130 Machine Learning and Neural Networks
Examination 22nd October 2014/Karhunen

(Voit vastata tenttiin myös suomeksi.)

1. Answer briefly (using a few lines) to the following questions or items:
   (a) What means whitening of the data?
   (b) Explain what is so-called bias-variance dilemma.
   (c) For what purpose is weight decay used?
   (d) What describes so-called U-matrix in context with self-organizing maps?
   (e) How are sub-Gaussian and super-Gaussian signals defined?
   (f) What kind of neural network structure is used in deep learning?

2. Assume that the relationship between the input vector $\mathbf{x}$ and the desired response (output) vector $\mathbf{d}$ is of the form
$$\mathbf{d} = \mathbf{h}(\mathbf{x}) + \mathbf{e}$$
where $\mathbf{h}(\mathbf{x})$ is the true mapping between $\mathbf{x}$ and $\mathbf{d}$ and $\mathbf{e}$ is the error or noise vector. Consider modeling the unknown true mapping $\mathbf{h}(\mathbf{x})$ by the output $\mathbf{y}(\mathbf{x}, \mathbf{w})$ of a neural network, where the vector $\mathbf{w}$ contains all the adjustable weights of the neural network. Assume that you have at your disposal N training pairs $(\mathbf{x}_i, \mathbf{d}_i)$ of the mapping. Show that if the training pairs are independent, and the noise vector $\mathbf{e}$ is Gaussian with zero mean and covariance matrix $\sigma^2\mathbf{I}$, the standard least-squares method and maximum likelihood method provide the same results.

3. Compare multilayer perceptron networks and support vector machines. Which general properties they have? What are their benefits and drawbacks when compared with each other?

4. Consider n zero mean source signals s_i, which have been collected to the vector $\mathbf{s} = (s_1, s_2, \ldots, s_n)^T$. Assume that the joint distribution of the sources is n variables multivariate Gaussian distribution
$$p(\mathbf{s}) = (2\pi)^{-n/2}(\det \mathbf{C})^{-1/2} \exp\left\{-\frac{1}{2}\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}\right\}$$
Here $\mathbf{C} = \mathrm{E}\{\mathbf{s}\mathbf{s}^T\}$ is the covariance matrix of the vector $\mathbf{s}$ and det means determinant. Assume that the source signals $s_1, s_2, \ldots, s_n$ are mutually uncorrelated. Show that they are mutually statistically independent.

   Does this propery hold for other than Gaussian distributed source signals? How about the reverse property: are statistically independent source signals source signals mutually uncorrelated?