

1. (Problem 8.15 Haykin) Let $\bar{k}_{ij}$ denote the centered counterpart of the ij -th element k_{ij} of the Gram $\mathbf{K}$. Derive the following formula (Schölkopf, 1977):

$$\bar{k}_{ij} = k_{ij} - \frac{1}{N} \sum_{m=1}^N \Phi^T(\mathbf{x}_m) \Phi(\mathbf{x}_j) - \frac{1}{N} \sum_{n=1}^N \Phi^T(\mathbf{x}_i) \Phi(\mathbf{x}_n) + \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N \Phi^T(\mathbf{x}_m) \Phi(\mathbf{x}_n) .$$

Suggest a compact representation of this relation in matrix form.

2. (Problem 8.16 Haykin) Show that the normalization of eigenvector α of the Gram $\mathbf{K}$ is equivalent to the requirement that Eq. (8.109) be satisfied.
3. (Problem 9.8 Haykin) Using the transformation formula of Eq. (9.40) applied to Eq. (9.37), derive the probability density function of Eq. (9.41). [Note: I believe Haykin's Eq. 9.41 is missing a $\frac{1}{\sigma}$ term. Please verify or disprove this.]
4. (Problem 9.9 Haykin) This problem is in two parts, addressing the issues involved in deriving a couple of equations that pertain to the kernel SOM algorithm:

(a) The *incomplete gamma distribution* of a random variable X , with sample value x , is defined by (Abramowitz and Stegun, 2965, p. 260):

$$P_X(x|\alpha) = \frac{1}{\Gamma(\alpha)} \int_0^x t^{\alpha-1} \exp(-t) dt$$

where $\Gamma(\alpha)$ is the gamma function. The *complement* of the incomplete gamma distribution is correspondingly defined by

$$\Gamma(\alpha, x) = \int_x^\infty t^{\alpha-1} \exp(-t) dt .$$

Using these two formulas, derive the cumulative distribution function of random variable R that is defined in Eq. (9.43). This propagates the error found in 9.8.

(b) Using the formula of the incomplete gamma distribution as the definition of the averaged neural output $\bar{y}_i$, derive Eq. (9.51) for the partial derivative $\partial \bar{y}_i(r) / \partial r$.

5. (Problem 9.10 Haykin) In developing the approximate update formula of Eq. (9.55) for the weight vector of the kernel SOM algorithm, we justified ignoring the second term in Eq. (9.52). Yet, in deriving the update formula of Eq. (9.58) for the kernel width σ_i , no approximation was made. Justify this latter choice.