T-61.5030 Advanced course in neural computing

Solutions for exercise 7

1. If we are not interested in the dependencies inside the groups $\mathbf{X}_1$ or $\mathbf{X}_2$ but are interested only on dependecies between these groups, we can use canonical correlation analysis (CCA). It means that we try to find projections from $\mathbf{X}_1$ and $\mathbf{X}_2$ which would be as correlated as possible. CCA uses only second-order statistics; for Gaussian variables this is sufficient but CCA can also be used for non-Gaussian variables. For a tutorial on CCA, see, e.g., the one by Magnus Borga available at http://people.imt.liu.se/~magnus/cca/ . The following is partly based on the information in that tutorial.

   The correlation coefficient between zero mean random variables a and b is defined to be $\rho_{ab} = E\{ab\}/\sqrt{E\{a^2\}E\{b^2\}}$. We have projections $a_i = \mathbf{u}_i^T\mathbf{x}_1$ and $b_i = \mathbf{v}_i^T\mathbf{x}_2$ and we would like to maximise the correlations between a_i and b_i. We restrict the solutions to be uncorrelated for different i: $E\{a_ia_j\} = E\{b_ib_j\} = E\{a_ib_j\} = 0$ for $i \neq j$.

   For the ith projections we have

$$\rho_{a_ib_i} = \frac{E\{\mathbf{u}_i^T\mathbf{x}_1\mathbf{x}_2^T\mathbf{v}_i\}}{\sqrt{E\{\mathbf{u}_i^T\mathbf{x}_1\mathbf{x}_1^T\mathbf{u}_i\}E\{\mathbf{v}_i^T\mathbf{x}_2\mathbf{x}_2^T\mathbf{v}_i\}}} = \frac{\mathbf{u}_i^T E\{\mathbf{x}_1\mathbf{x}_2^T\}\mathbf{v}_i}{\sqrt{\mathbf{u}_i^T E\{\mathbf{x}_1\mathbf{x}_1^T\}\mathbf{u}_i\mathbf{v}_i^T E\{\mathbf{x}_2\mathbf{x}_2^T\}\mathbf{v}_i}} = \frac{\mathbf{u}_i^T\mathbf{\Sigma}_{12}\mathbf{v}_i}{\sqrt{\mathbf{u}_i^T\mathbf{\Sigma}_1\mathbf{u}_i\mathbf{v}_i^T\mathbf{\Sigma}_2\mathbf{v}_i}} .$$

   It can be shown that the maximization corresponds to solving either one of the following eigenvalue equations:

$$\mathbf{\Sigma}_1^{-1}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_{12}^T\mathbf{u}_i = \rho_{a_ib_i}^2\mathbf{u}_i$$
$$\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_{12}^T\mathbf{\Sigma}_1^{-1}\mathbf{\Sigma}_{12}\mathbf{v}_i = \rho_{a_ib_i}^2\mathbf{v}_i .$$

   Connection to singular value decomposition. Canonical correlations are invariant to affine transformations (for example, if a transformation $\mathbf{A}$ is used for $\mathbf{x}_1$, just set $\hat{\mathbf{u}}_i = \mathbf{A}^{-1}\mathbf{u}_i$). Therefore, to simplify the situation, suppose that $\mathbf{X}_1$ and $\mathbf{X}_2$ have been whitened (see problem 4 for the precise transformations needed). Then $\mathbf{\Sigma}_1 = \mathbf{I}$ and $\mathbf{\Sigma}_2 = \mathbf{I}$. (Note that $\mathbf{X}_1$ and $\mathbf{X}_2$ can have different dimensionalities, so the two identity matrices can be of different sizes.) The eigenvalue equations then become

$$\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{12}^T\mathbf{u}_i = \rho_{a_ib_i}^2\mathbf{u}_i \ , \ \mathbf{\Sigma}_{12}^T\mathbf{\Sigma}_{12}\mathbf{v}_i = \rho_{a_ib_i}^2\mathbf{v}_i . \tag{1}$$

   Solving the above equations corresponds to *singular value decomposition* (SVD) of $\mathbf{\Sigma}_{12}$. SVD is similar to eigendecomposition but the orthogonal matrices need not be the same. This also means the SVD can be extended for non-square matrices. The singular value decomposition for matrix $\mathbf{\Sigma}_{12} = E\{\mathbf{x}_1\mathbf{x}_2^T\}$ gives a decomposition $\mathbf{\Sigma}_{12} = \mathbf{UDV}^T$, where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal square matrices and $\mathbf{D}$ is a matrix which has non-zero elements only on its diagonal. Notice that $\mathbf{D}$ has the same shape as $\mathbf{\Sigma}_{12}$ and it is therefore not necessarily square. The matrices $\mathbf{U}$ and $\mathbf{V}$ are computed by eigendecomposition:

$$\mathbf{\Sigma}_{12} \quad = \quad \mathbf{U}\mathbf{D}\mathbf{V}^T \quad \Rightarrow \quad \begin{cases} \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{12}^T = \mathbf{U}\mathbf{D}\mathbf{V}^T\mathbf{V}\mathbf{D}^T\mathbf{U}^T = \mathbf{U}\mathbf{D}\mathbf{D}^T\mathbf{U}^T \quad \text{and} \\ \mathbf{\Sigma}_{12}^T\mathbf{\Sigma}_{12} = \mathbf{V}\mathbf{D}^T\mathbf{U}^T\mathbf{U}\mathbf{D}\mathbf{V}^T = \mathbf{V}\mathbf{D}^T\mathbf{D}\mathbf{V}^T\,. \end{cases} \tag{2}$$

Therefore $\mathbf{U}$ and $\mathbf{V}$ are the same as the solutions to the eigenvalue equations (1): $\mathbf{U} = [\mathbf{u}_1 \dots \mathbf{u}_n]^T$ and $\mathbf{V} = [\mathbf{v}_1 \dots \mathbf{v}_m]^T$ and the diagonal elements of $\mathbf{D}$ are the corresponding correlation coefficients ρ_i.

Note: SVD can be seen as an extension to eigendecomposition. If SVD is done for the covariance matrix $\mathbf{\Sigma}_1 = E\{\mathbf{x}_1\mathbf{x}_1^T\}$, we have $\mathbf{\Sigma}_1\mathbf{\Sigma}_1^T = \mathbf{\Sigma}_1^T\mathbf{\Sigma}_1$ and $\mathbf{D}\mathbf{D}^T = \mathbf{D}^T\mathbf{D}$ in equation (2), and the orthogonal matrices $\mathbf{U}$ and $\mathbf{V}$ are therefore the same. The diagonal matrix then contains the eigenvalues which can be interpreted as variances to the directions given by the eigenvectors. For $\mathbf{\Sigma}_{12}$ the diagonal elements give the correlation coefficients (assuming that $\mathbf{\Sigma}_{12}$ is computed for the whitened data).

2. (a) We shall use subindices k and l instead of i and j in order to avoid confusion with the $i = \sqrt{-1}$. Denote the Fourier transform of the sequence $x_k(t)$ by $X_k(\omega)$ and the Fourier transform of $s_l(t)$ by $S_l(\omega)$. Recall that the Fourier transform changes a delay by D_{kl} into multiplication with term $e^{i\omega D_{kl}}$. Fourier transforming both sides of the equation defining the mixtures thus yields

$$X_k(\omega) = \sum_{l=1}^{N} a_{kl} e^{i\omega D_{kl}} S_l(\omega)\,.$$

This can be written as $\mathbf{X}(\omega) = \mathbf{A}(\omega)\mathbf{S}(\omega)$ when $\mathbf{X}$ and $\mathbf{S}$ are defined to be the vectors containing X_k and S_l and

$$\mathbf{A}(\omega) = \begin{pmatrix} a_{11}e^{i\omega D_{11}} & \cdots & a_{1N}e^{i\omega D_{1N}} \\ \vdots & \ddots & \vdots \\ a_{M1}e^{i\omega D_{M1}} & \cdots & a_{MN}e^{i\omega D_{MN}} \end{pmatrix}.$$

Notice that $\mathbf{A}$ is not constant.

(b) Since the original a_{kl} are real, we have $a_{kl} = \pm|A_{kl}|$, where A_{kl} denotes the element of matrix $\mathbf{A}$.

3. Assume that x_1 and x_2 are independent random variables. The kurtosis (fourth-order cumulant) of a random variable y is defined by

$$kurt(y) = E\{(y - E\{y\})^4\} - 3(E\{(y - E\{y\})^2\})^2.$$

Let us prove that $kurt(x_1 + x_2) = kurt(x_1) + kurt(x_2)$. We may without loss of generality assume that x_1 and x_2 are zero-mean. Then the kurtosis of $x_1 + x_2$ is

$$kurt(x_1 + x_2) = E\{(x_1 + x_2)^4\} - 3(E\{(x_1 + x_2)^2\})^2$$

$$= E\{x_1^4 + 4x_1^3x_2 + 6x_1^2x_2^2 + 4x_1x_2^3 + x_2^4\} - 3(E\{x_1^2 + 2x_1x_2 + x_2^2\})^2$$

Expectation is a linear operation, ie. $E\{\alpha y_1 + \beta y_2\} = \alpha E\{y_1\} + \beta E\{y_2\}$ for random variables y_1 and y_2 and scalar multipliers α and β. The above formula can therefore be rewritten as

$$\begin{aligned} kurt(x_1 + x_2) &= E\{x_1^4\} + 4E\{x_1^3 x_2\} + 6E\{x_1^2 x_2^2\} + 4E\{x_1 x_2^3\} + E\{x_2^4\} \\ &\quad -3(E\{x_1^2\} + 2E\{x_1 x_2\} + E\{x_2^2\})^2 \\ &= E\{x_1^4\} + 4E\{x_1^3 x_2\} + 6E\{x_1^2 x_2^2\} + 4E\{x_1 x_2^3\} + E\{x_2^4\} \\ &\quad -3E\{x_1^2\}^2 - 12E\{x_1^2\}E\{x_1 x_2\} - 6E\{x_1^2\}E\{x_2^2\} \\ &\quad -12E\{x_1 x_2\}^2 - 12E\{x_1 x_2\}E\{x_2^2\} - 3E\{x_2^2\}^2. \end{aligned}$$

Since x_1 and x_2 are independent, $E\{x_1^p x_2^q\} = E\{x_1^p\}E\{x_2^q\}$, for all $q, p \in \{1, \ldots, 4\}$. Then the above formula can be further rewritten:

$$\begin{aligned} kurt(x_1 + x_2) &= E\{x_1^4\} + 4E\{x_1^3\}E\{x_2\} + 6E\{x_1^2\}E\{x_2^2\} + 4E\{x_1\}E\{x_2^3\} + E\{x_2^4\} \\ &\quad -3E\{x_1^2\}^2 - 12E\{x_1^2\}E\{x_1\}E\{x_2\} - 6E\{x_1^2\}E\{x_2^2\} \\ &\quad -12E\{x_1\}^2 E\{x_2\}^2 - 12E\{x_1\}E\{x_2\}E\{x_2^2\} - 3E\{x_2^2\}^2. \end{aligned}$$

Here, $E\{x_1\} = E\{x_2\} = 0$, so the above reduces to

$$kurt(x_1 + x_2) = E\{x_1^4\} + 6E\{x_1^2\}E\{x_2^2\} + E\{x_2^4\} - 3E\{x_1^2\}^2 - 6E\{x_1^2\}E\{x_2^2\} - 3E\{x_2^2\}^2.$$

The terms $6E\{x_1^2\}E\{x_2^2\}$ and $-6E\{x_1^2\}E\{x_2^2\}$ cancel each other. Rearranging terms, we get

$$kurt(x_1 + x_2) = E\{x_1^4\} - 3E\{x_1^2\}^2 + E\{x_2^4\} - 3E\{x_2^2\}^2 = kurt(x_1) + kurt(x_2).$$

Let us prove the second property $kurt(\alpha x_1) = \alpha^4 kurt(x_1)$. We have

$$\begin{aligned} kurt(\alpha x_1) &= E\{(\alpha x_1)^4\} - 3(E\{(\alpha x_1)^2\})^2 = E\{\alpha^4 x_1^4\} - 3(E\{\alpha^2 x_1^2\})^2 \\ &= \alpha^4 E\{x_1^4\} - 3(\alpha^2 E\{x_1^2\})^2 = \alpha^4 E\{x_1^4\} - 3\alpha^4(E\{x_1^2\})^2 \\ &= \alpha^4(E\{x_1^4\} - 3(E\{x_1^2\})^2) = \alpha^4 kurt(x_1). \end{aligned}$$

4. Let $\mathbf{x}$ be the observed vector, and denote

$$\mathbf{x}_2 = \mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T\mathbf{x},$$

where $\mathbf{E}$ is the orthogonal matrix of eigenvectors of $E\{\mathbf{x}\mathbf{x}^T\}$, $\mathbf{D}$ is the diagonal matrix of its eigenvalues, $\mathbf{D} = \text{diag}(d_1, \ldots, d_n)$, and $\mathbf{D}^{-1/2}$ is a diagonal matrix whose diagonal elements are simply those of $\mathbf{D}$ raised to power $-1/2$, $\mathbf{D} = \text{diag}(d_1^{-1/2}, \ldots, d_n^{-1/2})$. Let us show that $\mathbf{x}_2$ is white.

Assume that $\mathbf{x}$ is zero-mean ($E\{\mathbf{x}\} = \mathbf{0}$). The matrices $\mathbf{E}$ and $\mathbf{D}$ are constant with regard to the expectation operator. Then

$$E\{\mathbf{x}_2\} = E\{\mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T\mathbf{x}\} = \mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T E\{\mathbf{x}\} = \mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T\mathbf{0} = \mathbf{0}.$$

It remains to show that $E\{\mathbf{x}_2\mathbf{x}_2^T\} = \mathbf{I}$. We have

$$E\{\mathbf{x}_2\mathbf{x}_2^T\} = E\{(\mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T\mathbf{x})(\mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T\mathbf{x})^T\} = E\{\mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T\mathbf{x}\mathbf{x}^T\mathbf{E}\mathbf{D}^{T-1/2}\mathbf{E}^T\}$$

$$= \mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T E\{\mathbf{x}\mathbf{x}^T\}\mathbf{E}\mathbf{D}^{T-1/2}\mathbf{E}^T,$$

where the last equality follows because $\mathbf{E}$ and $\mathbf{D}$ are constant with regard to expectation. From the definition of $\mathbf{E}$ and $\mathbf{D}$ we have the eigendecomposition

$$E\{\mathbf{x}\mathbf{x}^T\} = \mathbf{E}\mathbf{D}\mathbf{E}^T.$$

Inserting this into the previous equation we get

$$E\{\mathbf{x}_2\mathbf{x}_2^T\} = \mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T(\mathbf{E}\mathbf{D}\mathbf{E}^T)\mathbf{E}\mathbf{D}^{T-1/2}\mathbf{E}^T.$$

Since the eigenvector matrix $\mathbf{E}$ is orthogonal, we have $\mathbf{E}^T\mathbf{E} = \mathbf{E}\mathbf{E}^T = \mathbf{I}$. Applying this, we get

$$E\{\mathbf{x}_2\mathbf{x}_2^T\} = \mathbf{E}\mathbf{D}^{-1/2}(\mathbf{E}^T\mathbf{E})\mathbf{D}(\mathbf{E}^T\mathbf{E})\mathbf{D}^{T-1/2}\mathbf{E}^T$$

$$= \mathbf{E}\mathbf{D}^{-1/2}\mathbf{I}\mathbf{D}\mathbf{I}\mathbf{D}^{T-1/2}\mathbf{E}^T = \mathbf{E}(\mathbf{D}^{-1/2}\mathbf{D}\mathbf{D}^{T-1/2})\mathbf{E}^T$$

$$= \mathbf{E}(\operatorname{diag}(d_1^{-1/2},\ldots,d_n^{-1/2})\operatorname{diag}(d_1,\ldots,d_n)\operatorname{diag}(d_1^{-1/2},\ldots,d_n^{-1/2}))\mathbf{E}^T$$

$$= \mathbf{E}(\operatorname{diag}(d_1^{-1/2},\ldots,d_n^{-1/2})\operatorname{diag}(d_1^{1/2},\ldots,d_n^{1/2}))\mathbf{E}^T$$

$$= \mathbf{E}\mathbf{I}\mathbf{E}^T = \mathbf{E}\mathbf{E}^T = \mathbf{I}.$$

Therefore $\mathbf{x}_2$ is white.

5. Let us consider the distribution

$$g(x) = \frac{b}{4}\{\exp(-b|x-a|) + \exp(-b|x+a|)\}$$

where we assume $b > 0$ (otherwise $g(x)$ would not be a distribution).

(a) The n^{th} order moment of a distribution $p(x)$ with infinite support is

$$m_n = \int_{-\infty}^{\infty} p(x)x^n dx.$$

Note that $g(x)$ is symmetric about zero. Therefore the mean (first moment) of the distribution is zero (this can also be proved by integration). The kurtosis of $g(x)$ (the kurtosis of a random variable x distributed according to $g(x)$) is then defined by

$$kurt_{g(x)}(x) = E\{x^4\} - 3(E\{x^2\})^2 = m_4 - 3m_2^2.$$

Let us calculate m_4 and m_2. For m_n we have

$$m_n = \int_{-\infty}^{\infty} g(x)x^n dx = \frac{b}{4}\int_{-\infty}^{\infty}\{\exp(-b|x-a|) + \exp(-b|x+a|)\}x^n dx$$

$$= \frac{b}{4}\int_{-\infty}^{\infty}\exp(-b|x-a|)x^n dx + \frac{b}{4}\int_{-\infty}^{\infty}\exp(-b|x+a|)x^n dx$$

$$= \frac{b}{4}\left[\int_{-\infty}^{a} e^{b(x-a)}x^n dx + \int_{a}^{\infty} e^{-b(x-a)}x^n dx + \int_{-\infty}^{-a} e^{b(x+a)}x^n dx + \int_{-a}^{\infty} e^{-b(x+a)}x^n dx\right]$$

$$= \frac{b}{4}\left[e^{-ba}\int_{-\infty}^{a} e^{bx}x^n dx + e^{ba}\int_{a}^{\infty} e^{-bx}x^n dx + e^{ba}\int_{-\infty}^{-a} e^{bx}x^n dx + e^{-ba}\int_{-a}^{\infty} e^{-bx}x^n dx\right]$$

In our case, n is even, so x^n is symmetric about zero. Let us then change the integration variable from x to $-x$ in the second and fourth integrals. We get

$$m_n = \frac{b}{4}\left[e^{-ba}\int_{-\infty}^{a} e^{bx}x^n dx - e^{ba}\int_{-a}^{-\infty} e^{bx}x^n dx + e^{ba}\int_{-\infty}^{-a} e^{bx}x^n dx - e^{-ba}\int_{a}^{-\infty} e^{bx}x^n dx\right]$$

$$= \frac{b}{4}\left[e^{-ba}\int_{-\infty}^{a} e^{bx}x^n dx + e^{ba}\int_{-\infty}^{-a} e^{bx}x^n dx + e^{ba}\int_{-\infty}^{-a} e^{bx}x^n dx + e^{-ba}\int_{-\infty}^{a} e^{bx}x^n dx\right]$$

$$= \frac{b}{2}\left[e^{-ba}\int_{-\infty}^{a} e^{bx}x^n dx + e^{ba}\int_{-\infty}^{-a} e^{bx}x^n dx\right].$$

The remaining integrals can be computed by partial integration. For $n = 4$ we have

$$\int_{-\infty}^{c} e^{bx}x^4 dx = \frac{c^4}{b}e^{bc} - \int_{-\infty}^{c} \frac{4}{b}e^{bx}x^3 dx$$

$$= \left[\frac{c^4}{b} - \frac{4c^3}{b^2}\right]e^{bc} + \int_{-\infty}^{c} \frac{12}{b^2}e^{bx}x^2 dx = \left[\frac{c^4}{b} - \frac{4c^3}{b^2} + \frac{12c^2}{b^3}\right]e^{bc} - \int_{-\infty}^{c} \frac{24}{b^3}e^{bx}x dx$$

$$= \left[\frac{c^4}{b} - \frac{4c^3}{b^2} + \frac{12c^2}{b^3} - \frac{24c}{b^4}\right]e^{bc} + \int_{-\infty}^{c} \frac{24}{b^4}e^{bx} dx = \left[\frac{c^4}{b} - \frac{4c^3}{b^2} + \frac{12c^2}{b^3} - \frac{24c}{b^4} + \frac{24}{b^5}\right]e^{bc}$$

Inserting this result into the formula for m_4, with a and $-a$ in place of c, respectively, we get

$$m_4 = \frac{b}{2}e^{-ba+ba}\left[\left(\frac{a^4}{b} - \frac{4a^3}{b^2} + \frac{12a^2}{b^3} - \frac{24a}{b^4} + \frac{24}{b^5}\right) + \left(\frac{a^4}{b} + \frac{4a^3}{b^2} + \frac{12a^2}{b^3} + \frac{24a}{b^4} + \frac{24}{b^5}\right)\right]$$

$$= b\left[\frac{a^4}{b} + \frac{12a^2}{b^3} + \frac{24}{b^5}\right].$$

Performing a similar partial integration procedure for m_2, we get

$$m_2 = \frac{b}{2}\left[e^{-ba}e^{ba}\left(\frac{a^2}{b} - \frac{2a}{b^2} + \frac{2}{b^3}\right) + e^{ba}e^{-ba}\left(\frac{a^2}{b} + \frac{2a}{b^2} + \frac{2}{b^3}\right)\right] = b\left[\frac{a^2}{b} + \frac{2}{b^3}\right].$$

Inserting these results into the moment-based formula for the kurtosis, we have

$$kurt_{g(x)}(x) = m_4 - 3m_2^2 = b\left[\frac{a^4}{b} + \frac{12a^2}{b^3} + \frac{24}{b^5}\right] - 3b^2\left[\frac{a^4}{b^2} + \frac{4a^2}{b^4} + \frac{4}{b^6}\right] = \frac{12 - 2a^4b^4}{b^4}.$$

Let us assume that the distribution $g(x)$ has unit variance ($m_2 = 1$). Then

$$b\left[\frac{a^2}{b} + \frac{2}{b^3}\right] = a^2 + \frac{2}{b^2} = 1 \Rightarrow (a^2 + \frac{2}{b^2})^2 = a^4 + \frac{4a^2}{b^2} + \frac{4}{b^4} = 1$$

$$\Rightarrow b^4 = b^4a^4 + 4a^2b^2 + 4 \Rightarrow kurt_{g(x)}(x) = \frac{12 - 2a^4b^4}{4 + 4a^2b^2 + a^4b^4}.$$

Here we require $b^2 > 0 \Rightarrow a^2 < 1$, where the the right-hand side follows from the unit variance assumption. In this case, the previous formula is in fact simpler.

(b) Since $b^4 > 0$, the sign of the kurtosis depends only on the numerator term of the above expression.

i. When the kurtosis is negative, we have

$$12 - 2a^4b^4 < 0 \Rightarrow a^4 > \frac{6}{b^4} \Rightarrow a^2 > \frac{\sqrt{6}}{b^2} \Rightarrow |a| > \frac{\sqrt[4]{6}}{b}.$$

With the unit variance assumption, we have $b^4 = 4/(1 - 2a^2 + a^4)$, from which

$$a^4 > \frac{6}{b^4} = \frac{3}{2}(1 - 2a^2 + a^4) \Rightarrow a^4 - 6a^2 + 3 < 0, a^2 > 0$$

$$\Rightarrow 3-\sqrt{6} < a^2 < 3+\sqrt{6} \Rightarrow \sqrt{3 - \sqrt{6}} < |a| < \sqrt{3 + \sqrt{6}} \Rightarrow \sqrt{3 - \sqrt{6}} < |a| < 1\ .$$

The corresponding values for b can be calculated from the unit variance assumption (see the formula above).

ii. The kurtosis is zero when $|a| = \sqrt[4]{6}/b$. In the unit variance case this becomes $|a| = \sqrt{3 - \sqrt{6}}$ (2 points).

iii. The kurtosis is positive when $|a| < \sqrt[4]{6}/b$. In the unit variance case this becomes $|a| < \sqrt{3 - \sqrt{6}}$ (1 interval).

(c) Distributions (or the associated random variables) with negative kurtosis are called subgaussian and distributions with positive kurtosis are called supergaussian. Here, the distribution $g(x)$ can be either, depending on the values of a and b (see the previous section). In the figure below, the solid curve is subgaussian, the dashed curve has zero kurtosis and the dashdot curve is supergaussian.