DD2434 Practical 3

$$N(x|\mu,\Sigma) = \frac{1}{(2\pi)^{\frac{D}{2}}|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}\underbrace{(x-\mu)^T\Sigma^{-1}(x-\mu)}}$$

$(x-\mu)^T\Sigma^{-1}(x-\mu)$

1. Σ = diagonal matrix

$$\Sigma = \begin{bmatrix} \Sigma_1 & & \emptyset \\ & \Sigma_2 & \\ \emptyset & & \ddots \; \Sigma_d \end{bmatrix} \Rightarrow \Sigma^{-1} = \begin{bmatrix} \frac{1}{\Sigma_1} & & \emptyset \\ & \frac{1}{\Sigma_2} & \\ \emptyset & & \ddots \; \frac{1}{\Sigma_d} \end{bmatrix}$$

$$(x-\mu)^T\Sigma^{-1}(x-\mu) = [(x_1-\mu_1), \ldots, (x_D-\mu_D)] \begin{bmatrix} \frac{1}{\Sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\Sigma_D} \end{bmatrix} \begin{bmatrix} x_1-\mu_1 \\ \vdots \\ x_D-\mu_D \end{bmatrix} =$$

$$= (x_1-\mu_1)\cdot\frac{1}{\Sigma_1}\cdot(x_1-\mu_1) + \ldots + (x_D-\mu_D)\cdot\frac{1}{\Sigma_D}(x_D-\mu_D) =$$

$$= \sum_i^D \underbrace{\frac{1}{\Sigma_i}(x_i-\mu_i)^2}_{A}$$

A - alway positive
- small value $(x_i-\mu_i)$ → close to mean
- Σ_i - scales this value

⇒ Σ_i - big ⇒ uncertain about this dimenson → large deviation from mean doesn't matter

Σ_i - small ⇒ certain about this dimension ⇒ large deviation from mean matters a lot

2 Σ - square matrix

eigenvalue decomposition

$$\Sigma = U \Lambda U^T$$

$$\Sigma^{-1} = (U \Lambda U^T)^{-1} = U^{-T} \Lambda^{-1} U^{-1} = \{U^T U = I\} =$$

$$= U \Lambda^{-1} U^T$$

$$(x-\mu)^T \Sigma^{-1} (x-\mu) = (x-\mu)^T U \Lambda^{-1} U^T (x-\mu) =$$

$$= (x-\mu)^T U \Lambda^{-\frac{1}{2}} \Lambda^{-\frac{1}{2}} U^T (x-\mu) =$$

$$= \{(AB)^T = B^T A^T\} = \underbrace{(U^T \Lambda^{-\frac{1}{2}} (x-\mu))}_{A}{}^T (U^T \Lambda^{-\frac{1}{2}} (x-\mu))$$

$$A = \underbrace{U^T \Lambda^{-\frac{1}{2}}}\,(x-\mu)$$

↳ This is just a linear mapping of the deviation from the mean.

<u>view 1</u>

$$(U^T \Lambda^{-\frac{1}{2}} (x-\mu))^T \, I \, (U^T \Lambda^{-\frac{1}{2}} (x-\mu))$$

- The mapping $U^T \Lambda^{-\frac{1}{2}}$ maps the data to a representation to a space where the covariance is spherical

<u>view 2</u>

$$(U^T (x-\mu))^T \Lambda^{-1} (U^T (x-\mu))$$

- The mapping U^T maps the data to a representation to a space where the covariance is diagonal.

IMPORTANT: The transformation makes the dimensions independent.

Posterior Distribution

General normal:

$$N(\mu,\Sigma) \propto e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)} =$$

$$= e^{-\frac{1}{2}x^T\Sigma^{-1}x} \cdot e^{x^T\Sigma^{-1}\mu} \cdot \underbrace{e^{-\frac{1}{2}\mu^T\Sigma^{-1}\mu}}_{\text{const.}} \qquad (*)$$

$$p(y|w,x) = N(wx, \sigma^2 I)$$

$$p(w) = N(0, \Sigma^{-1})$$

$$p(w|y,x) \propto p(y|w,x)p(w) \propto$$

$$\propto e^{-\frac{1}{2\sigma^2}(y-Xw)^T(y-Xw)} \cdot e^{-\frac{1}{2}w^T\Sigma^{-1}w}$$

Lets look at the exponent

$$\Rightarrow -\frac{1}{2\sigma^2}(y-Xw)^T(y-Xw) - \frac{1}{2}w^T\Sigma^{-1}w = \{(*)\} =$$

$$= \underbrace{-\frac{1}{2\sigma^2}y^Ty}_{A} + \underbrace{\frac{1}{\sigma^2}y^T(Xw)}_{B} \underbrace{- \frac{1}{2\sigma^2}(Xw)^T(Xw) - \frac{1}{2}w^T\Sigma^{-1}w}_{C}$$

Identify: A - our new constant term

B - our new mixed term

C - the new term with quadratic in parameters.

$$C: -\frac{1}{2\sigma^2}(Xw)^T(Xw) - \frac{1}{2}w^T\Sigma^{-1}w =$$

$$= \left\{(AB)^T = B^TA^T\right\} = -\frac{1}{2\sigma^2}w^TX^TXw - \frac{1}{2}w^T\Sigma^{-1}w =$$

$$= -\frac{1}{2}w^T\left(\frac{1}{\sigma^2}X^TX\right)w - \frac{1}{2}w^T\Sigma^{-1}w =$$

$$= -\frac{1}{2}w^T\left(\frac{1}{\sigma^2}X^TX + \Sigma^{-1}\right)w$$

Identify: $S^{-1} = \frac{1}{\sigma^2}X^TX + \Sigma^{-1}$

$$B: \frac{1}{\sigma^2}y^T(Xw) = \left\{(AB)^T = B^TA^T\right\} = \frac{1}{\sigma^2}w^TX^Ty$$

Where does the mean pop up?

- In the mixed term

$$x^T\Sigma^{-1}\mu \Rightarrow w^TS^{-1}\mu = w^T\left(\frac{1}{\sigma^2}X^TX + \Sigma^{-1}\right)\mu = \frac{1}{\sigma^2}w^TX^Ty$$

Solve for μ:

$$\cancel{w^T}\left(\frac{1}{\sigma^2}X^TX + \Sigma^{-1}\right)\mu = \frac{1}{\sigma^2}\cancel{w^T}X^Ty$$

$$\left(\frac{1}{\sigma^2}X^TX + \Sigma^{-1}\right)\mu = \frac{1}{\sigma^2}X^Ty$$

$$\Rightarrow \mu = \frac{1}{\sigma^2}\left(\frac{1}{\sigma^2}X^TX + \Sigma^{-1}\right)^{-1}X^Ty$$

$$p(w|Y,X) \propto N\left(\frac{1}{\sigma^2}\left(\frac{1}{\sigma^2}X^TX + \Sigma^{-1}\right)^{-1}X^Ty,\ \frac{1}{\sigma^2}X^TX + \Sigma^{-1}\right)$$

This procedure is called "Completing the Square"

Conditional Distribution

$$p(x_1, x_2) = N\left(\begin{bmatrix}\mu_1 \\ \mu_2\end{bmatrix}, \begin{bmatrix}\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22}\end{bmatrix}\right)$$

Goal: want to find $p(x_1|x_2)$

$$p(x_1, x_2) = p(x_1|x_2)\,p(x_2)$$

$$p(x_2) = N(\mu_2, \Sigma_{22})$$

$$p(x_1, x_2) \propto \exp\left(-\frac{1}{2}\begin{pmatrix}x_1 - \mu_1 \\ x_2 - \mu_2\end{pmatrix}^T \begin{pmatrix}\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}^{-1} \begin{pmatrix}x_1 - \mu \\ x_2 - \mu\end{pmatrix}\right)$$

① we want to factor out

$$p(x_2) \propto \exp\left(-\frac{1}{2}(x_2 - \mu_2)^T \Sigma_{22}^{-1} (x_2 - \mu_2)\right)$$

from this exponent

② - If we can re-write the covariance in such a manner so that it factorises, i.e. becomes block diagonal.

- How do we invert the covariance matrix to keep Σ_{22} isolated?

Schur Complement

$$M = \begin{bmatrix} E & F \\ G & H \end{bmatrix} \qquad \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}^{-1} = \begin{bmatrix} A_1^{-1} & 0 \\ 0 & A_2^{-1} \end{bmatrix}$$

① Want to find the inverse of M.

$$\begin{bmatrix} I & A \\ 0 & I \end{bmatrix} \begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} E+AG & F+AH \\ G & H \end{bmatrix}$$

② want clear out $F+AH$

$$\Rightarrow F+AH=0 \Rightarrow AH=-F \Rightarrow A=-FH^{-1}$$

$$\Rightarrow \begin{bmatrix} I & -FH^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} E-FH^{-1}G & 0 \\ G & H \end{bmatrix}$$

$$\begin{bmatrix} E-FH^{-1}G & 0 \\ G & H \end{bmatrix} \begin{bmatrix} I & 0 \\ B & I \end{bmatrix} = \begin{bmatrix} E-FH^{-1}G & 0 \\ G+HB & H \end{bmatrix}$$

③ Want to clear out

$$\Rightarrow G-HB=0 \Rightarrow HB=-G \Rightarrow B=-H^{-1}G$$

$$\begin{bmatrix} E-FH^{-1}G & 0 \\ G & H \end{bmatrix} \begin{bmatrix} I & 0 \\ -H^{-1}G & I \end{bmatrix} = \begin{bmatrix} E-FH^{-1}G & 0 \\ 0 & H \end{bmatrix}$$

Put it all together

$$\underbrace{\begin{bmatrix} I & -FH^{-1} \\ 0 & I \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} E & F \\ G & H \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} I & 0 \\ -H^{-1}G & I \end{bmatrix}}_{B} = \underbrace{\begin{bmatrix} E-FH^{-1}G & 0 \\ 0 & H \end{bmatrix}}_{C}$$

$$(AMB)^{-1} = C^{-1}$$

$$(AMB)^{-1} = B^{-1}M^{-1}A^{-1}$$

$$B^{-1}M^{-1}A^{-1} = C^{-1}$$

$$\underbrace{BB^{-1}}_{I}M^{-1}A^{-1} = BC^{-1} \Rightarrow M^{-1}A^{-1} = BC^{-1}$$

$$M^{-1}A^{-1}A = BC^{-1}A \Rightarrow M^{-1} = BC^{-1}A$$

$$M^{-1} = \begin{bmatrix} I & 0 \\ -HG^{-1} & I \end{bmatrix} \begin{bmatrix} (E-FH^{-1}G)^{-1} & 0 \\ 0 & H^{-1} \end{bmatrix} \begin{bmatrix} I & -FH^{-1} \\ 0 & I \end{bmatrix} =$$

$$= \begin{bmatrix} (E-FH^{-1}G)^{-1} & -(E-FH^{-1}G)^{-1}FH^{-1} \\ -H^{-1}G(E-FH^{-1}G)^{-1} & H^{-1}+H^{-1}G(E-FH^{-1}G)^{-1}FH^{-1} \end{bmatrix}$$

Now we have our inverse co-varians.

The term $E-FH^{-1}G$ is the Schur complement of M with respect to H (M/H).

Look at exponent. ⑨

$$-\frac{1}{2}\begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix}^T \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix} =$$

$$= -\frac{1}{2}\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}^T \begin{bmatrix} I & 0 \\ -\Sigma_{22}^{-1}\Sigma_{21} & I \end{bmatrix} \begin{bmatrix} (\Sigma/\Sigma_{22})^{-1} & 0 \\ 0 & \Sigma_{22}^{-1} \end{bmatrix} \begin{bmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix} =$$

$$= -\frac{1}{2}\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}^T \begin{bmatrix} I & 0 \\ -\Sigma_{22}^{-1}\Sigma_{21} & I \end{bmatrix} \begin{bmatrix} (\Sigma/\Sigma_{22})^{-1} & -(\Sigma/\Sigma_{22})^{-1}\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & \Sigma_{22}^{-1} \end{bmatrix} \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix} =$$

$$= -\frac{1}{2}\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}^T \begin{bmatrix} (\Sigma/\Sigma_{22})^{-1} & -(\Sigma/\Sigma_{22})^{-1}\Sigma_{12}\Sigma_{22}^{-1} \\ -\Sigma_{21}\Sigma_{22}^{-1}(\Sigma/\Sigma_{22})^{-1} & \Sigma_{22}^{-1} \end{bmatrix} \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}$$

$$= -\frac{1}{2}\left(x_1-\mu_1-\Sigma_{21}\Sigma_{22}^{-1}(x_2-\mu_2)\right)^T(\Sigma/\Sigma_{22})^{-1}\left(x-\mu_1-\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu)\right) -$$

$$\underbrace{-\frac{1}{2}(x_2-\mu_2)^T\Sigma_{22}^{-1}(x_2-\mu_2)}_{p(x_2)}$$

$$\Rightarrow p(x_1|x_2) \propto \exp\left(-\frac{1}{2}(x_1-\mu_1-\Sigma_{21}\Sigma_{22}^{-1}(x_2-\mu_2))^T(\Sigma/\Sigma_{22})^{-1}(x-\mu_1-\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu))\right)$$

$\Rightarrow$ mean: $\mu_{1|2} = \mu_1 + \Sigma_{21}\Sigma_{22}^{-1}(x_2-\mu_2)$

variance: $\Sigma_{1|2} = \Sigma/\Sigma_{22} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$

1. What happens if x_1 & x_2 are independent?

$$\Sigma_{12} = \emptyset$$

$$\mu_{1|2} = \mu_1 + \emptyset \cdot \Sigma_{22}^{-1}(x_2 - \mu_2) = 0$$

$$\Sigma_{1|2} = \Sigma_{11} - 0\,\Sigma_{22}^{-1}\,0 = \Sigma_{11}$$

2. If they are completely co-dependent

$$\Sigma_{12} = \Sigma_{22} = \Sigma_{11}$$

$$\mu_{1|2} = \mu_1 + \underbrace{\Sigma_{22} \cdot \Sigma_{22}^{-1}}_{I}(x_2 - \mu_2) = x_2 + (\mu_1 - \mu_2)$$

$$\Sigma_{1|2} = \Sigma_{11} - \Sigma_{22}\Sigma_{22}^{-1}\Sigma_{22} = \Sigma_{11} - \Sigma_{22} = 0$$

3. μ: The term $\Sigma_{12}\Sigma_{22}^{-1}$ adjusts the mean based on the co-variation

Σ: $\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ adjusts the variance based on how similar the cross-covariance is to the variations in x_2.

Gaussian Marginals

$$p(x_1, x_2) = N\left(\begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right)$$

Precision Matrix $\Lambda = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1}$

$$p(x_1) = \int p(x_1, x_2)\, dx_2$$

① Lets look at the exponent of the joint.

$$-\frac{1}{2}(x_1-\mu_1)^T\Lambda_{11}(x_1-\mu_1) - \frac{1}{2}(x_1-\mu_1)^T\Lambda_{12}(x_2-\mu_2) -$$
$$-\frac{1}{2}(x_2-\mu_2)^T\Lambda_{21}(x_1-\mu_1) - \frac{1}{2}(x_2-\mu_2)^T\Lambda^{22}(x_2-\mu_2) =$$

$$= -\frac{1}{2}\left(x_2^T\Lambda_{22}x_2 - 2x_2^T\Lambda_{22}(\mu_2 - \Lambda_{22}^{-1}\Lambda_{21}(x_1-\mu_1)) - 2x_1^T\Lambda_{12}\mu_2 + \right.$$
$$\left. + 2\mu_1^T\Lambda_{12}\mu_2 + \mu_2^T\Lambda_{22}\mu_2 + x_1^T\Lambda_{11}x_1 - 2x_1^T\Lambda_{11}\mu_1 + \mu_1^T\Lambda_{11}\mu_1\right)$$

② We know the marginal is Gaussian so lets try to "complete the square".
$\Rightarrow$ Put terms with same precision together.

$$-\frac{1}{2}(x_2-(\mu_2-\Lambda_{22}^{-1}\Lambda_{21}(x_1-\mu_1)))^T\Lambda_{22}(x_2-(\mu_2-\Lambda_{22}^{-1}\Lambda_{21}(x_1-\mu_1)))+$$

$$* \quad +\frac{1}{2}(x_1^T\Lambda_{12}\Lambda_{22}^{-1}\Lambda_{21}x_1-2x_1^T\Lambda_{12}\Lambda_{22}^{-1}\Lambda_{12}\mu_1+\mu_1^T\Lambda_{12}\Lambda_{22}^{-1}\Lambda_{21}\mu_1)-$$

$$** \quad -\frac{1}{2}(x_1^T\Lambda_{11}x_1-2x_1^T\Lambda_{11}\mu_1+\mu_1^T\Lambda_{11}\mu_1)=$$

3 Complete the square for the last two terms.

$$* \quad \frac{1}{2}(x_1^T\Lambda_{12}\Lambda_{22}^{-1}\Lambda_{21}x_1-2x_1^T\Lambda_{12}\Lambda_{22}^{-1}\Lambda_{12}\mu_1+\mu_1^T\Lambda_{12}\Lambda_{22}^{-1}\Lambda_{21}\mu_1)=$$

$$=\{\Lambda_{12}=\Lambda_{21}\}=$$

$$=\frac{1}{2}((x_1-\mu_1)^T(\Lambda_{12}\Lambda_{22}^{-1}\Lambda_{21})(x_1-\mu_1))$$

$$** \quad \frac{1}{2}(x_1^T\Lambda_{11}x_1-2x_1^T\Lambda_{11}\mu_1+\mu_1^T\Lambda_{11}\mu_1)=$$

$$=\frac{1}{2}((x_1-\mu_1)^T\Lambda_{11}(x_1-\mu_1))$$

Put everything together

$$-\frac{1}{2}(x_2-(\mu_2-\Lambda_{22}^{-1}\Lambda_{21}(x_1-\mu_1)))^T\Lambda_{22}(x_2-(\mu_2-\Lambda_{22}^{-1}\Lambda_{21}(x_1-\mu_1)))-$$

$$-\frac{1}{2}(x_1-\mu_1)(\Lambda_{11}-\Lambda_{12}\Lambda_{22}^{-1}\Lambda_{21})(x_1-\mu_1)$$

We have two exponents

$$p(x_1, x_2) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma|^{\frac{1}{2}}} \cdot \exp(E_1) \cdot \exp(E_2)$$

$$E_1 = -\frac{1}{2}\left(x_2 - (\mu_2 - \Lambda_{22}^{-1}\Lambda_{21}(x_1 - \mu_1))\right)^T \Lambda_{22}\left(x_2 - (\mu_2 - \Lambda_{22}^{-1}\Lambda_{21}(x_1 - \mu_1))\right)$$

$$E_2 = -\frac{1}{2}(x_1 - \mu_1)^T \left(\Lambda_{11} - \Lambda_{12}\Lambda_{22}^{-1}\Lambda_{21}\right)(x_1 - \mu_1)$$

$$p(x_1) = \int p(x_1, x_2)\,dx_2 = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma|^{\frac{1}{2}}} \int \exp(E_1)\exp(E_2)\,dx_2 =$$

$$= \left\{ E_2 \text{ is independent of } x_2 \right\} =$$

$$= \frac{1}{(2\pi)^{\frac{N}{2}} |\Sigma|^{\frac{1}{2}}} \int \exp(E_1)\,dx_2 \exp(E_2) =$$

= a density function always integrates to one.

$$\Rightarrow \int \exp(E_1)\,dx_2 =$$

$$\int p(y)\,dy = 1$$

$$p(y) = \frac{1}{(2\pi)^{\frac{d}{2}} |S|^{\frac{1}{2}}} \cdot \int e^{-\frac{1}{2}(y-\mu_y)^T S^{-1}(y-\mu_y)}\,dy = 1$$

$$\Rightarrow \int \exp(E_1)\,dx = (2\pi)^{\frac{D_2}{2}} \cdot |\Lambda_{22}^{-1}|^{\frac{1}{2}}$$

$$\Rightarrow p(x_1) = \underbrace{(2\pi)^{\frac{D_2}{2}} \cdot |\Lambda_{22}^{-1}|^{\frac{1}{2}} \cdot \frac{1}{(2\pi)^{\frac{D}{2}} \cdot |\Sigma|^{\frac{1}{2}}}}_{\frac{1}{(2\pi)^{\frac{D-D_2}{2}} \cdot |\Lambda_{22}^{-1}|^{\frac{1}{2}} |\Sigma|^{\frac{1}{2}}}} \cdot \exp(E_2) =$$

Determinant

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A| \, |D - CA^{-1}B|$$

$$\Rightarrow |\Sigma| = |\Sigma_{11}| \, |\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}|$$

Schur complement

$$\Rightarrow \Lambda_{22}^{-1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$$

$$|\Lambda_{22}^{-1}|^{\frac{1}{2}} \cdot |\Sigma|^{\frac{1}{2}} = |\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}|^{-\frac{1}{2}} |\Sigma_{11}|^{\frac{1}{2}} |\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}|^{\frac{1}{2}} =$$

$$= |\Sigma_{11}|^{\frac{1}{2}}$$

$$\Rightarrow p(x_1) = \left\{ D = D_1 + D_2 \right\} =$$

$$= \frac{1}{(2\pi)^{\frac{D_1}{2}} |\Sigma_{11}|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}(x_1 - \mu_1)^T (\Lambda_{11} - \Lambda_{12}\Lambda_{22}^{-1}\Lambda_{21})(x_1 - \mu_1)} =$$

$$= \left\{ \Sigma_{11}^{-1} = \Lambda_{11} - \Lambda_{12}\Lambda_{22}^{-1}\Lambda_{21} \right\} =$$

$$= \frac{1}{(2\pi)^{\frac{D_1}{2}} |\Sigma_{11}|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}(x_1 - \mu_1)\Sigma_{11}^{-1}(x_1 - \mu_1)}$$

$$\underline{\underline{= N(\mu_1, \Sigma_{11})}}$$