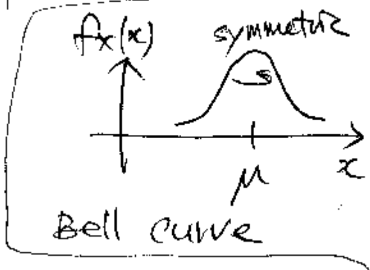
Real-valued Gaussian Random Vectors

○ A Gaussian random variable

- Def. A random variable X is Gaussian if its probability density function (pdf) $f_X(x)$ is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$



- A pdf can be understood as a limit of the histogram

- Properties of a Gaussian r.v.

(i) Its pdf is parameterized only by two constants μ and σ (or equivalently σ^2)

(ii) It turns out that $E[X] = \mu$ and $Var[X] = \sigma^2$.

(iii) Often denoted by $X \sim \mathcal{N}(\mu, \sigma^2)$

'normal' means Gaussian

(iv) $$E\{(X-\mu)^n\} = \begin{cases} 0, & \text{for odd } n \\ \dfrac{n!}{(\frac{n}{2})!}\left(\dfrac{\sigma}{\sqrt{2}}\right)^n, & \text{for even } n. \end{cases}$$

/ex/
$$E\{(X-\mu)^2\} = 1 \cdot \sigma^2$$
$$E\{(X-\mu)^4\} = 1 \cdot 3 \cdot \sigma^4$$
$$E\{(X-\mu)^6\} = 1 \cdot 3 \cdot 5\ \sigma^6$$
$$\vdots$$

(v) Its characteristic function is given by

$$\Phi_X(\omega) = \exp\left(j\omega\mu - \frac{1}{2}\omega^2\sigma^2\right)$$

It is the Fourier transform of $f_X(x)$ with sign for ω reversed. If $\mu = 0$, $\Phi_X(\omega)$ is also a bell curve.

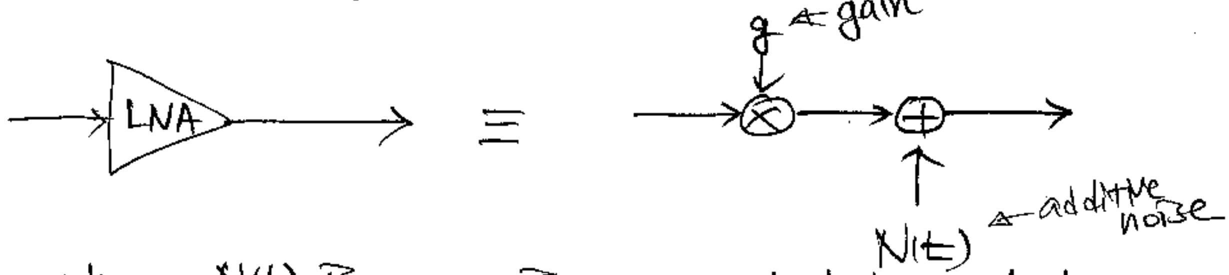
- Central Limit Theorem (CLT)

  If X_i's are independent and identically distributed (i.i.d.) random variables with $E[X_i] = \mu$ and $Var[X_i] = \sigma^2$, then

$$Y_n \triangleq \frac{(X_1-\mu)+(X_2-\mu)+\cdots+(X_n-\mu)}{\sigma\sqrt{n}} \xrightarrow{D} Y \sim N(0,1)$$

, i.e., Y_n converges to a zero-mean unit-variance Gaussian random variable in distribution.

In other words, the cdf of Y_n is very close to that of zero-mean unit-variance Gaussian if n is sufficiently large.

- 

where $N(t)$ is a noise generated by electrons that bounce back and forth while they flow. Since there are a lot of electrons, we can apply the CLT to claim that we observe a Gaussian noise at each time instant.

Wozencraft & Jacob "Principles of Communication Engineering"

variance of a uniformly distributed random variable is $b^2/12$, where b is the width of the density function. Specifying the variance essentially constrains the effective width of the density function. Figure 2.40 illustrates this effect for the Gaussian density function.

A precise statement of the constraint is due to Chebyshev. Let y be a

Bell(shaped) Curve

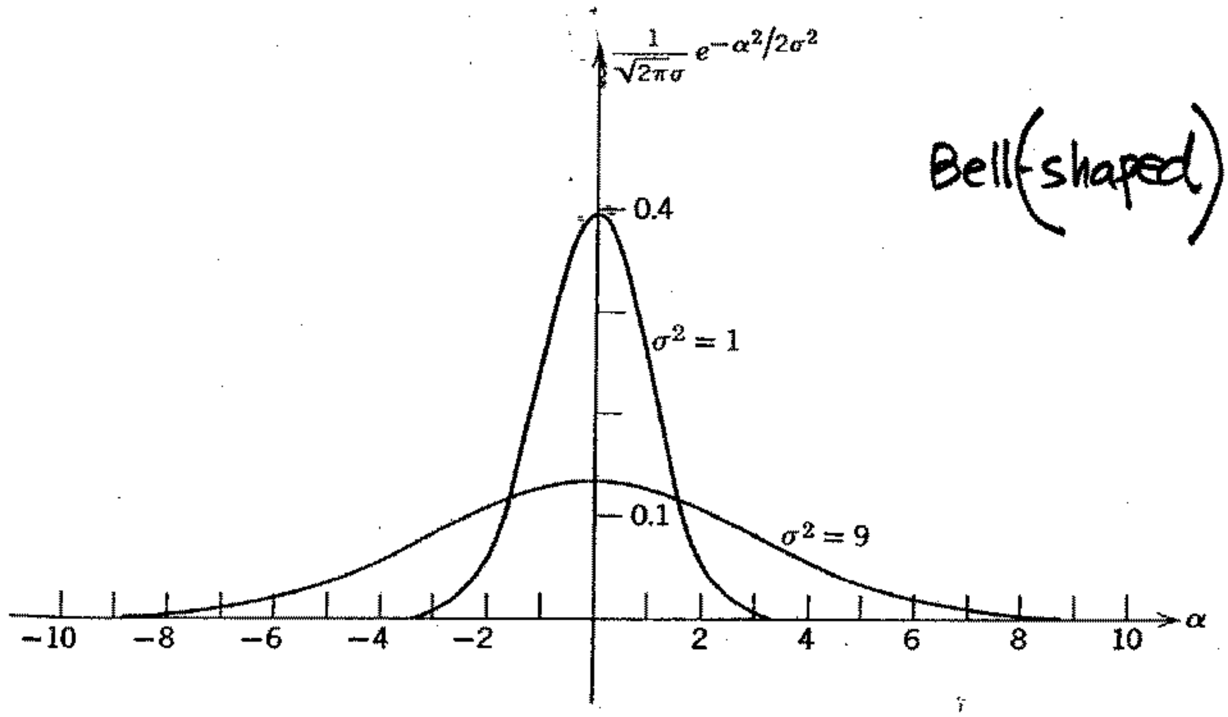


Figure 2.40 The Gaussian probability density function for two values of variance.

zero-mean random variable with finite variance σ_y^2. *Chebyshev's inequality* states that for any positive number ϵ

$$P[|y| \geqslant \epsilon] \leqslant \frac{{\sigma_y}^2}{\epsilon^2}; \qquad \bar{y} = 0. \tag{2.146}$$

Equation 2.146 can be proved as follows. By definition,

$$\overline{y^2} = \int_{-\infty}^{\infty} \alpha^2\, p_y(\alpha)\, d\alpha.$$

Since the integrand is positive,

$$\overline{y^2} \geqslant \int_{|\alpha| \geqslant \epsilon} \alpha^2\, p_y(\alpha)\, d\alpha.$$

This bound can be weakened further by replacing α^2 with its smallest value, ϵ^2, which yields

○ A Gaussian Random Vector

· Motivation.

Let's consider an affine transform of $\underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}$,

where X_i's are i.i.d. zero-mean unit-variance Gaussian random variables, i.e.

$$\underline{Y} = A\underline{X} + \underline{b}$$

Then, it can be shown that the pdf of $\underline{Y}$ is given by

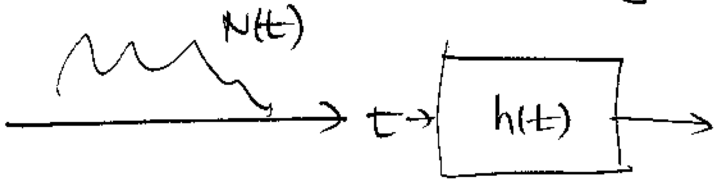
$$f_{\underline{Y}}(\underline{y}) = \frac{1}{\sqrt{2\pi}^N \sqrt{\det AA^T}} \exp\left(-\frac{1}{2}(\underline{y}-\underline{b})^T (AA^T)^{-1} (\underline{y}-\underline{b})\right)$$

when AA^T is invertible.

Since $E[\underline{Y}] = \underline{b}$ and $Cov[\underline{Y}] \triangleq E\{(\underline{Y}-E(\underline{Y}))(\underline{Y}-E(\underline{Y}))^T\}$
$= AA^T$

we may generalize this pdf and define a Gaussian random vector.

· The reason why an affine transform of independent Gaussian random variables is important is that if we filter the noise generated in LNA



we need to linearly combining independent Gaussian random variables. Here, we still want the resulting random variables called Gaussian.

$= \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}$

- Def. A random vector $\underline{X}$ is called Gaussian if its pdf is given by

$$f_{\underline{X}}(\underline{x}) = \frac{1}{\sqrt{2\pi}^N \sqrt{\det C}} \exp\left(-\frac{1}{2}(\underline{x}-\underline{\mu})^T C^{-1} (\underline{x}-\underline{\mu})\right)$$

for some $\underline{\mu}$ and $C = C^T \underset{\sim}{>} 0$
(This means positive definiteness of C)

- A pdf of a random vector can be understood as a limit of the multi-dimensional histogram.

- Properties of a Gaussian random vector.

(i) Its pdf is parameterized only by a vector μ and a positive definite symmetric matrix C.

(ii) It turns out that $E[\underline{X}] = \underline{\mu}$, $Cov[\underline{X}] = C$.

(iii) Often denoted by $\underline{X} \sim N(\underline{\mu}, C)$

(iv) Often called a <u>multivariate</u> Gaussian random variable.
↳ uni-variate

(v) Its characteristic function is given by

$$\Phi_{\underline{X}}(\underline{\omega}) = \exp\left(j\underline{\omega}^T \underline{\mu} - \frac{1}{2}\underline{\omega}^T C \underline{\omega}\right)$$

where $C = C^T \underset{\sim}{\geq} 0$

C can be positive semi-definite.!

→ This makes the definition using the characteristic function more powerful than that using the pdf.

(vi) If C^{-1} does not exist, the pdf cannot be written in terms of elementary functions. Instead, multi-dimensional Dirac delta functions must be used.

→ We encounter impulse fences.

(vii) If $\underline{X}$ is Gaussian, then $X_1, X_2, \cdots, X_N$ are called jointly Gaussian (random variables.)

(viii) Theorem.

Any affine transform $\underline{Y} = A\underline{X} + \underline{b}$ of $\underline{X} \sim N(\mu, C)$ is Gaussian.

/proof/

$$\underline{\Phi}_{\underline{Y}}(\underline{\omega}) \triangleq E[e^{j\underline{\omega}^T\underline{Y}}] = E[e^{j\underline{\omega}^T(A\underline{X}+\underline{b})}]$$

$$= e^{j\underline{\omega}^T\underline{b}}\, E[e^{j(A^T\underline{\omega})^T\underline{X}}] = e^{j\underline{\omega}^T\underline{b}}\, \underline{\Phi}_{\underline{X}}(A^T\underline{\omega})$$

$$= e^{j\underline{\omega}^T\underline{b}} \exp\left(j(A^T\underline{\omega})^T\underline{\mu} - \frac{1}{2}(A^T\underline{\omega})^T C (A^T\underline{\omega})\right)$$

$$= \exp\left(j\underline{\omega}^T(A\underline{\mu}+\underline{b}) - \frac{1}{2}\underline{\omega}^T(ACA^T)\underline{\omega}\right)$$

which is the characteristic function of a Gaussian random vector with mean $A\underline{\mu}+\underline{b}$ and covariance ACA^T

(ix) When $N=2$, $\underline{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ $C = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$, $\rho \neq \pm 1$

$$\Rightarrow f_{\underline{X}}(\underline{x}) = f_{X_1X_2}(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)} \left\{ \left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 \right\} \right]$$

(x) For $N=2$, $\underline{X} \sim N(\mu_1, \mu_2 ; \sigma_1, \sigma_2, \rho)$ or

$\underline{X} \sim N(\mu_1, \mu_2 ; \sigma_1^2, \sigma_2^2, \rho)$

bi-variate Gaussian

s shown in Fig. 2.26, this densit function may be visualized as tw "fences" of impulses at $\alpha_1 = 1$ and $\alpha_2 = 2$.

For a simple example of the use of a joint density function to calculate a probability, consider the event A defined by

$$A = \{\omega: x_1^2(\omega) + x_2^2(\omega) < c^2\}$$

and the two-dimensional Gaussian density function of Eq. 2.58, with the

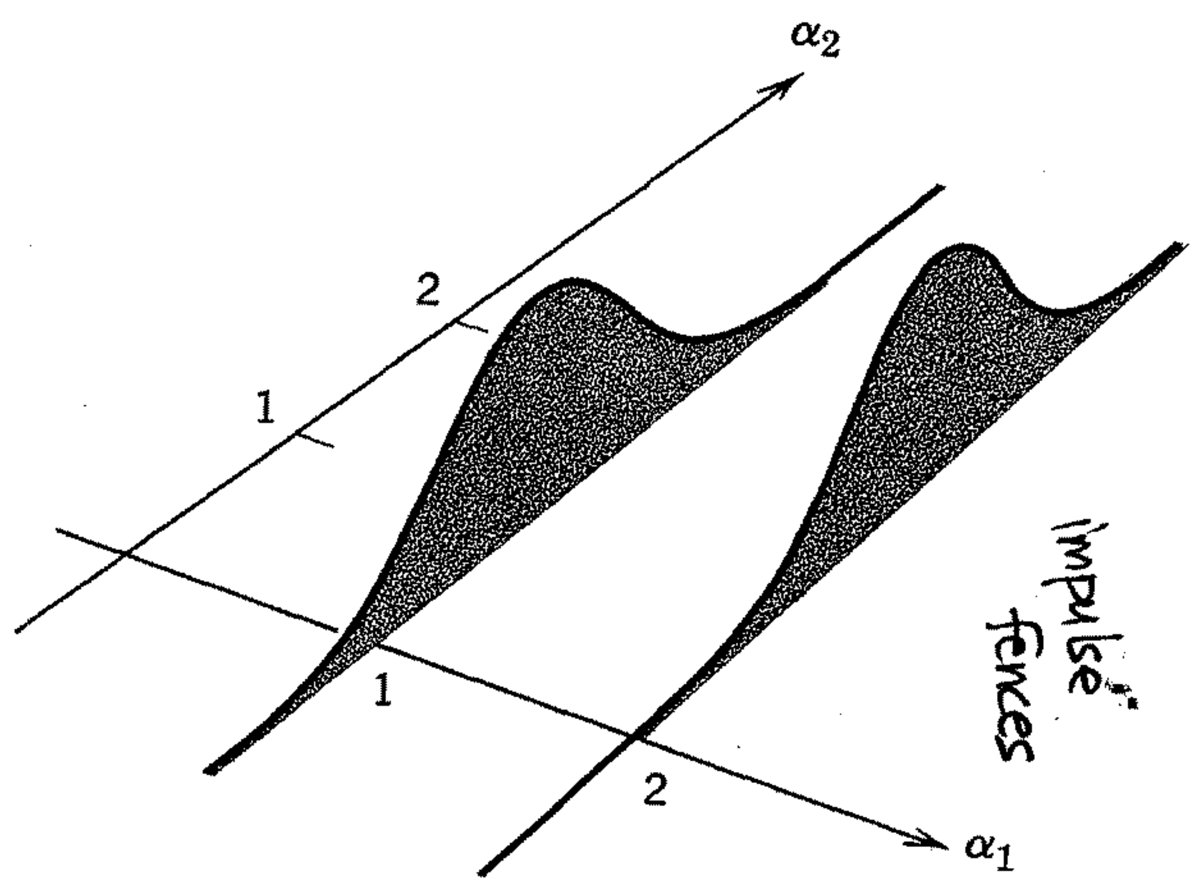


Figure 2.26 Two "fences" of impulses. The value of the one-dimensional impulse at $\alpha_1 = 1$ (or $\alpha_1 = 2$) depends on α_2.

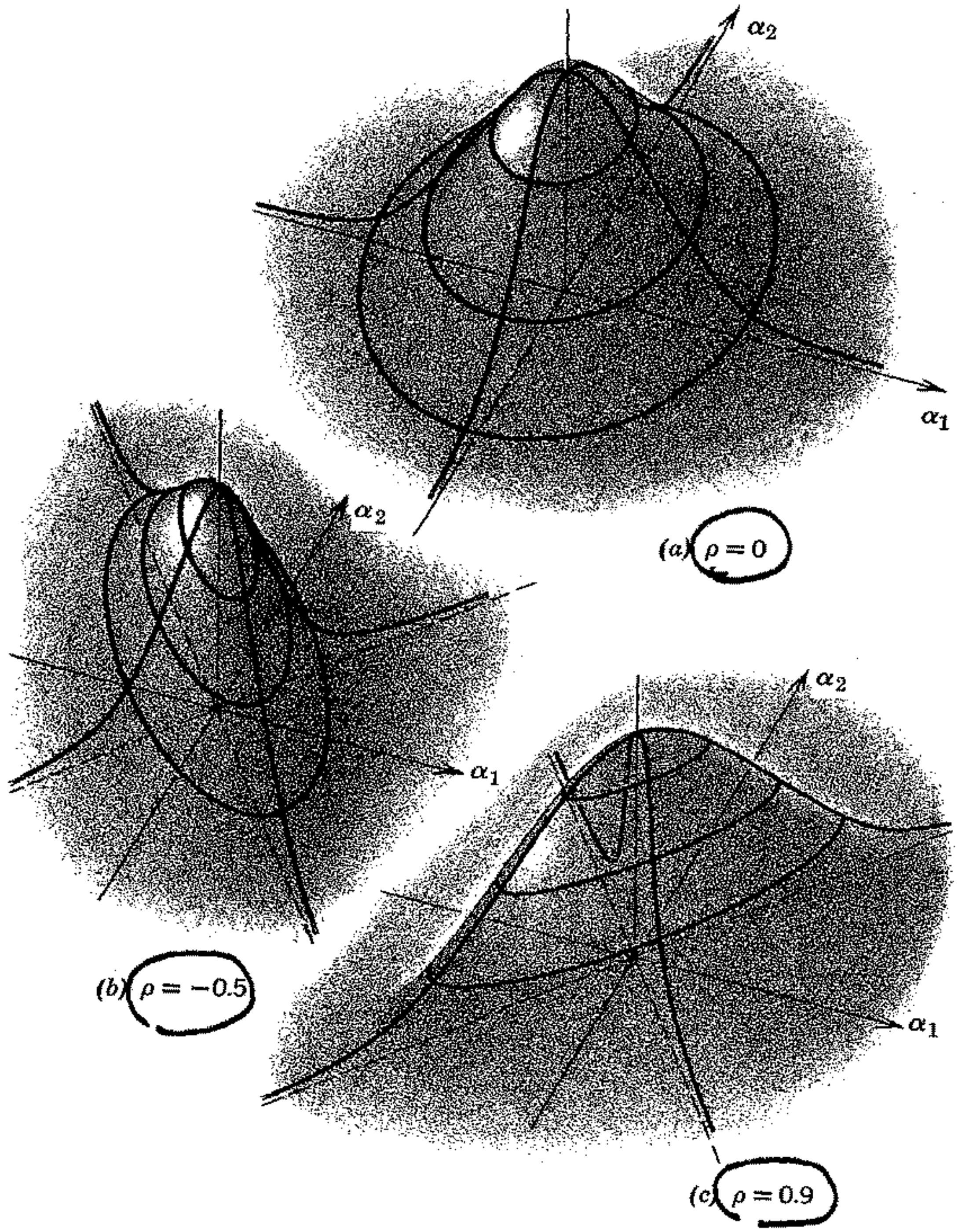


Figure 2.24 Examples of the two-dimensional Gaussian density function.

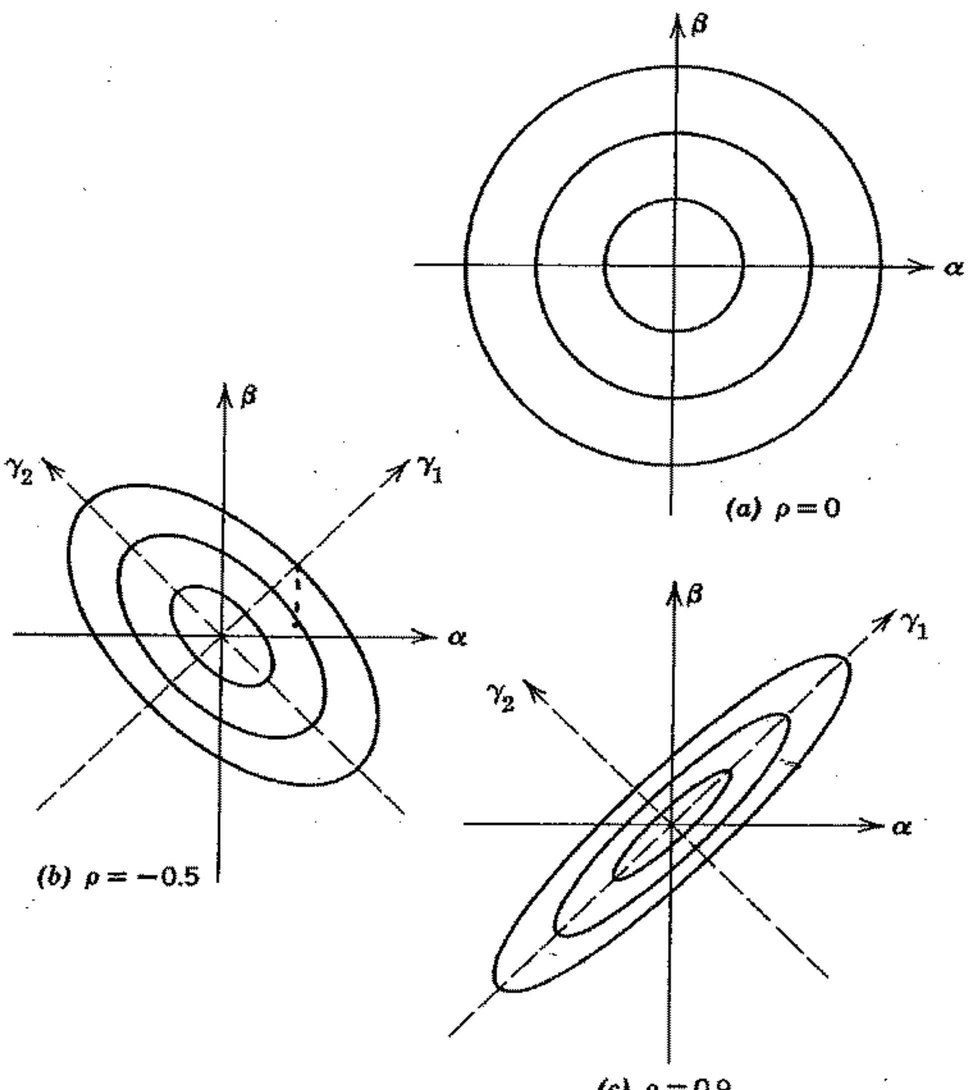


Figure 3.17 Contour plots of constant probability density for the two-dimensional Gaussian density function of Eq. 3.31. The density functions themselves are illustrated in Fig. 2.24 for $\sigma^2 = 1$.

Further insight into the behavior of p_{n_1,n_2} as a function of ρ can be gained from the contour plots of constant probability density shown in Fig. 3.17. The contours are most easily visualized in terms of coordinates γ_1, γ_2 rotated 45° from α, β. If we let

$$\alpha = \gamma_1 \cos\frac{\pi}{4} - \gamma_2 \sin\frac{\pi}{4}, \tag{3.35a}$$

$$\beta = \gamma_1 \sin\frac{\pi}{4} + \gamma_2 \cos\frac{\pi}{4}, \tag{3.35b}$$