Computer Vision

Simon J.D. Prince (2012)

p. 40

3.6 Normal-scaled inverse gamma distribution

The *normal-scaled inverse gamma distribution* (figure 3.6) is defined over a pair of continuous values μ, σ^2, the first of which can take any value and the second of which is constrained to be positive. As such it can define a distribution over the mean and variance parameters of the normal distribution.

The normal-scaled inverse gamma has four parameters $\alpha, \beta, \gamma, \delta$ where α, β, and γ are positive real numbers but δ can take any value. It has pdf:

$$Pr(\mu, \sigma^2) = \frac{\sqrt{\gamma}}{\sigma\sqrt{2\pi}} \frac{\beta^\alpha}{\Gamma[\alpha]} \left(\frac{1}{\sigma^2}\right)^{\alpha+1} \exp\left[-\frac{2\beta + \gamma(\delta - \mu)^2}{2\sigma^2}\right], \tag{3.13}$$

or for short

$$Pr(\mu, \sigma^2) = \text{NormInvGam}_{\mu,\sigma^2}[\alpha, \beta, \gamma, \delta]. \tag{3.14}$$

正規分布（normal distribution）に現れるパラメーター、平均（mean）μ と分散（variance）σ^2 に関する事前分布（prior）として、よく用いられる。

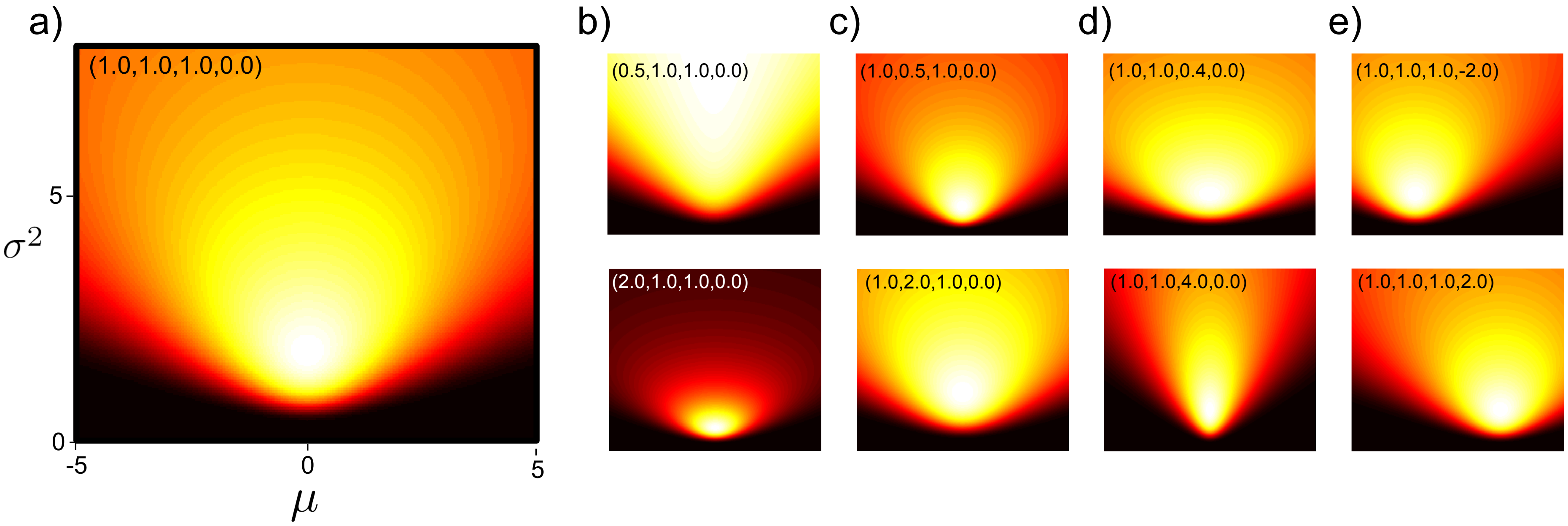

Figure 3.6 The normal-scaled inverse gamma distribution defines a probability distribution over bivariate continuous values μ, σ^2 where $\mu \in [-\infty, \infty]$ and $\sigma^2 \in [0, \infty]$. a) Distribution with parameters $[\alpha, \beta, \gamma, \delta] = [1, 1, 1, 0]$. b) Varying α. c) Varying β. d) Varying γ. e) Varying δ.

Problem 3.8 Calculate an expression for the mode (position of the peak in μ, σ^2 space) of the normal scaled inverse gamma distribution in terms of the parameters $\alpha, \beta, \gamma, \delta$.

$$
\begin{aligned}
&\text{NormInvGam}_{\mu,\sigma^2}[\alpha,\beta,\gamma,\delta] \\
&\qquad = \frac{\sqrt{\gamma}}{\sigma\sqrt{2\pi}}\,\frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\sigma^2}\right)^{\alpha+1}\exp\left[-\frac{2\beta+\gamma(\delta-\mu)^2}{2\sigma^2}\right], \\
&\qquad = \underbrace{\sqrt{\frac{\gamma}{2\pi\sigma^2}}\exp\left[-\frac{\gamma(\delta-\mu)^2}{2\sigma^2}\right]}_{\text{Norm}}\;\underbrace{\frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\sigma^2}\right)^{\alpha+1}\exp\left[-\frac{\beta}{\sigma^2}\right]}_{\text{InvGam}}.
\end{aligned}
$$

$$
\begin{aligned}
&\int_{-\infty}^{\infty} d\mu\,\text{NormInvGam}_{\mu,\sigma^2}[\alpha,\beta,\gamma,\delta] \\
&\qquad = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-x^2}dx\;\left[\frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\sigma^2}\right)^{\alpha+1}e^{-\beta/\sigma^2}\right], \quad \text{where } x\equiv\sqrt{\frac{\gamma}{2\sigma^2}}(\mu-\delta) \\
&\qquad = \frac{1}{\Gamma(\alpha)}\left(\frac{\beta}{\sigma^2}\right)^{\alpha+1}e^{-\beta/\sigma^2}\;\frac{1}{\beta}, \\
&\qquad \equiv \text{InvGam}_{\sigma^2}[\alpha,\beta].
\end{aligned}
$$

$$
\begin{aligned}
\int_0^\infty d\sigma^2 \int_{-\infty}^\infty d\mu \, \mathrm{NormInvGam}_{\mu,\sigma^2}[\alpha,\beta,\gamma,\delta] \\
&= \int_0^\infty \mathrm{InvGam}_{\sigma^2}[\alpha,\beta] \; d\sigma^2, \\
&= \int \mathrm{InvGam}_{\sigma^2}[\alpha,\beta] \left| \frac{d\sigma^2}{d\lambda} \right| \; d\lambda.
\end{aligned}
$$

$\lambda = 1/\sigma^2$ (precision) に選ぶと、

$$
\begin{aligned}
\mathrm{InvGam}_{\sigma^2}[\alpha,\beta] \left| \frac{d\sigma^2}{d\lambda} \right| &= \frac{1}{\Gamma(\alpha)} (\beta\lambda)^{\alpha-1} e^{-\beta\lambda} \beta, \\
&\equiv \mathrm{Gam}(\lambda|\alpha,\beta) \qquad \text{(after Bishop)}.
\end{aligned}
$$

$$
\begin{aligned}
\int_0^\infty d\sigma^2 \int_{-\infty}^\infty d\mu \, \mathrm{NormInvGam}_{\mu,\sigma^2}[\alpha,\beta,\gamma,\delta] &= \int_0^\infty \mathrm{Gam}(\lambda|\alpha,\beta) \; d\lambda, \\
&= \frac{1}{\Gamma(\alpha)} \int_0^\infty X^{\alpha-1} e^{-X} dX, \quad \text{(where } \; X \equiv \beta\lambda) \\
&= \frac{1}{\Gamma(\alpha)} \Gamma(\alpha), \\
&= 1.
\end{aligned}
$$

Summary:

$$
\begin{aligned}
&\mathrm{NormInvGam}_{\mu,\sigma^2}[\alpha,\beta,\gamma,\delta] \\
&\qquad = \frac{\sqrt{\gamma}}{\sigma\sqrt{2\pi}}\frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\sigma^2}\right)^{\alpha+1}\exp\left[-\frac{2\beta+\gamma(\delta-\mu)^2}{2\sigma^2}\right], \\
&\qquad = \underbrace{\sqrt{\frac{\gamma}{2\pi\sigma^2}}\exp\left[-\frac{\gamma(\delta-\mu)^2}{2\sigma^2}\right]}_{\mathrm{Norm}}\;\underbrace{\frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\sigma^2}\right)^{\alpha+1}\exp\left[-\frac{\beta}{\sigma^2}\right]}_{\mathrm{InvGam}}.
\end{aligned}
$$

Transformation: $(\mu,\sigma^2) \quad \to \quad (x, X),$

$$
\text{where} \qquad x = \sqrt{\frac{\gamma}{2\sigma^2}}(\mu-\delta), \qquad X = \frac{\beta}{\sigma^2}.
$$

Then

$$
\left|\frac{\partial(x\,X)}{\partial(\mu\,\sigma^2)}\right| = \frac{\partial x}{\partial \mu}\left|\frac{dX}{d\sigma^2}\right| = \sqrt{\frac{\gamma}{2\sigma^2}}\frac{X^2}{\beta}.
$$

$$\mathrm{NormInvGam}_{\mu,\sigma^2}[\alpha,\beta,\gamma,\delta] = \frac{1}{\sqrt{\pi}}e^{-x^2}\,\frac{1}{\Gamma(\alpha)}X^{\alpha-1}e^{-X}\left|\frac{\partial(x\,X)}{\partial(\mu\,\sigma^2)}\right|.$$

$$\begin{aligned}
\int_0^{\infty} d\sigma^2 \int_{-\infty}^{\infty} d\mu\, \mathrm{NormInvGam}_{\mu,\sigma^2}[\alpha,\beta,\gamma,\delta]& \\
&= \int\int \frac{1}{\sqrt{\pi}}e^{-x^2}\,\frac{1}{\Gamma(\alpha)}X^{\alpha-1}e^{-X}\left|\frac{\partial(x\,X)}{\partial(\mu\,\sigma^2)}\right| d\mu d\sigma^2, \\
&= \int\int \frac{1}{\sqrt{\pi}}e^{-x^2}\,\frac{1}{\Gamma(\alpha)}X^{\alpha-1}e^{-X}\,dxdX, \\
&= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-x^2}dx \;\; \frac{1}{\Gamma(\alpha)}\int_0^{\infty} X^{\alpha-1}e^{-X}dX, \\
&= \frac{1}{\sqrt{\pi}}\sqrt{\pi} \;\; \frac{1}{\Gamma(\alpha)}\Gamma(\alpha), \\
&= 1.
\end{aligned}$$

Return to

Problem 3.8 Calculate an expression for the mode (position of the peak in μ, σ^2 space) of the normal scaled inverse gamma distribution in terms of the parameters $\alpha, \beta, \gamma, \delta$.

$$\text{NormInvGam}_{\mu,\sigma^2}[\alpha,\beta,\gamma,\delta] \propto \left(\frac{1}{\sigma^2}\right)^{\alpha+3/2} \exp\left[-\frac{2\beta+\gamma(\delta-\mu)^2}{2\sigma^2}\right].$$

$$\ln\left(\text{NormInvGam}\right) = -\left(\alpha+\frac{3}{2}\right)\ln\sigma^2 - \frac{2\beta+\gamma(\delta-\mu)^2}{2\sigma^2} + \text{constant}.$$

$$\frac{\partial\,\text{NormInvGam}}{\partial\mu} = \frac{\gamma(\delta-\mu)}{\sigma^2}\,\text{NormInvGam}, \\ = 0,$$

$$\therefore \quad \mu = \delta. \tag{a}$$

$$\frac{\partial\,\text{NormInvGam}}{\partial\sigma^2} = \left[-\left(\alpha+\frac{3}{2}\right)\frac{1}{\sigma^2} + \frac{2\beta+\gamma(\delta-\mu)^2}{2\sigma^4}\right]\text{NIG},$$

$$= \frac{1}{2\sigma^4}\left[2\beta + \gamma(\delta-\mu)^2 - (2\alpha+3)\sigma^2\right] \text{NIG},$$

$$= 0,$$

$$\therefore \quad \sigma^2 = \frac{2\beta + \gamma(\delta-\mu)^2}{2\alpha+3}. \tag{b}$$

From Eqs. (a) and (b),

$$(\mu, \sigma^2) = \left(\delta, \frac{2\beta}{2\alpha+3}\right).$$

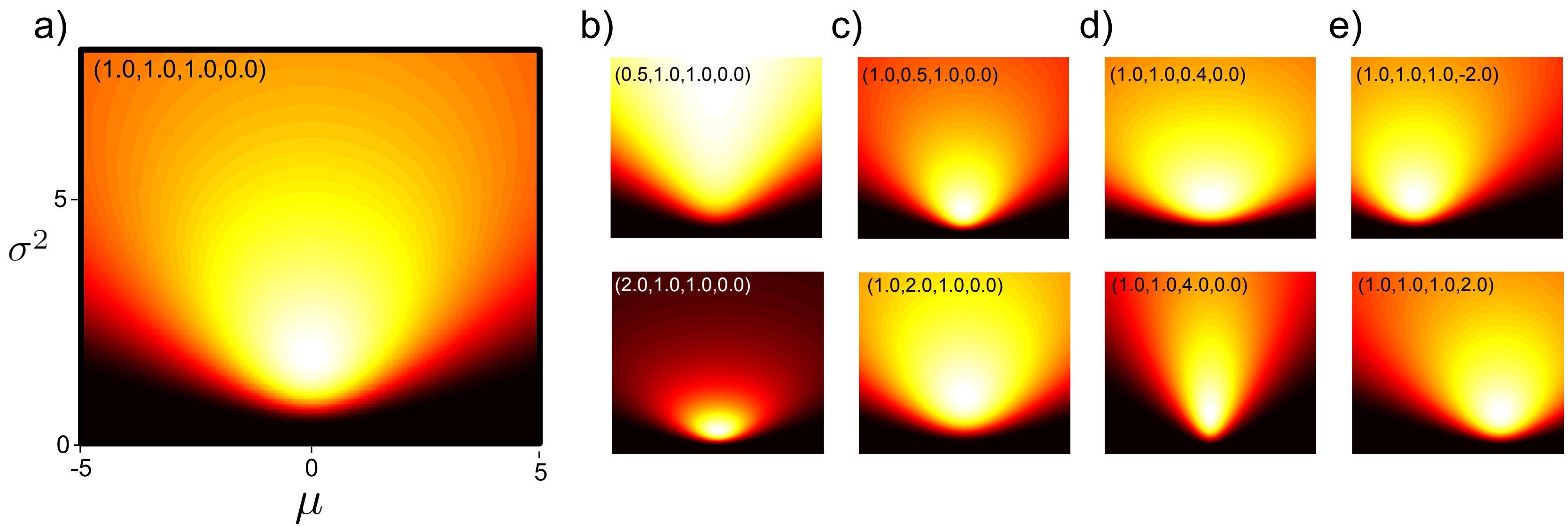

Figure 3.6 The normal-scaled inverse gamma distribution defines a probability distribution over bivariate continuous values μ, σ^2 where $\mu \in [-\infty, \infty]$ and $\sigma^2 \in [0, \infty]$. a) Distribution with parameters $[\alpha, \beta, \gamma, \delta] = [1, 1, 1, 0]$. b) Varying α. c) Varying β. d) Varying γ. e) Varying δ.