Change of variables formula: multivariate version

(This fixes some typos in Section 11.5.5 of my book.)
We now extend the results of Section 11.4.5 to joint densities.
Let g be a function that maps $\mathbb{R}^n$ to $\mathbb{R}^m$, and let $\mathbf{y} = g(\mathbf{x})$. Define the **Jacobian matrix** as

$$\mathbf{J}_{\mathbf{x}\to\mathbf{y}} \stackrel{\text{def}}{=} \frac{\partial(y_1,\ldots,y_m)}{\partial(x_1,\ldots,x_n)} \stackrel{\text{def}}{=} \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix} \tag{1}$$

If $m = n$, then $\mathbf{J}$ is square, and we can compute its determinant; denote that by $J = \det \mathbf{J}$. Also, let $|J|$ be the sign of the determinant. This measures how much a unit volume changes when we apply g.
If g is an invertible mapping, we can define the pdf of the transformed variables in terms of the original variables as follows:

$$p_y(\mathbf{y}) = p_x(\mathbf{x})|\det\left(\frac{\partial \mathbf{x}}{\partial \mathbf{y}}\right)| = p_x(\mathbf{x})|\det \mathbf{J}_{\mathbf{y}\to\mathbf{x}}| = p_x(\mathbf{x})|J_{\mathbf{y}\to\mathbf{x}}| \tag{2}$$

As an example, consider transforming a density from **Cartesian** coordinates $\mathbf{x} = (x_1, x_2)$ to **polar** coordinates $\mathbf{y} = (r, \theta)$, where $x_1 = r\cos\theta$ and $x_2 = r\sin\theta$. Then

$$\mathbf{J}_{\mathbf{y}\to\mathbf{x}} = \begin{pmatrix} \frac{\partial x_1}{\partial r} & \frac{\partial x_1}{\partial \theta} \\ \frac{\partial x_2}{\partial r} & \frac{\partial x_2}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \tag{3}$$

and

$$|\det \mathbf{J}| = |J| = |r\cos^2\theta + r\sin^2\theta| = |r| \tag{4}$$

Hence

$$p_{\mathbf{y}}(\mathbf{y}) = p_{\mathbf{x}}(\mathbf{x})|J| \tag{5}$$

$$p_{R,\Theta}(r,\theta) = p_{X_1,X_2}(x_1,x_2)r \tag{6}$$

To see this geometrically, notice that

$$P(r \le R \le r + dr, \theta \le \Theta \le \theta + d\theta) = p_{R,\Theta}(r,\theta)drd\theta \tag{7}$$

is the area of the shaded patch in Figure 1, which is clearly $rdrd\theta$, times the density at the center of the patch. Hence

$$P(r \le R \le r + dr, \theta \le \Theta \le \theta + d\theta) = p_{R,\Theta}(r,\theta)drd\theta \tag{8}$$

$$= p_{X,Y}(r\cos\theta, r\sin\theta)r\ dr\ d\theta \tag{9}$$

Hence

$$p_{R,\Theta}(r,\theta) = p_{X,Y}(r\cos\theta, r\sin\theta)r \tag{10}$$

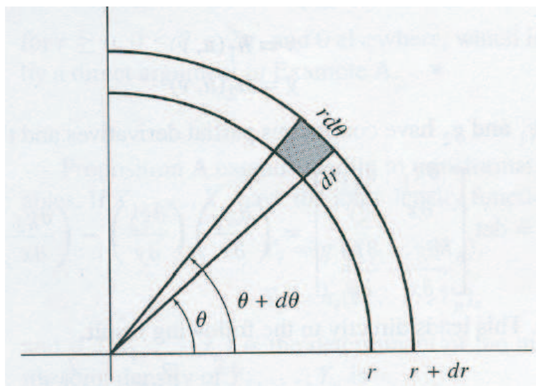

Figure 1: Change of variables from polar to Cartesian. The area of the shaded patch is $rdrd\theta$. Source: **?** Figure 3.16.