In later chapters it will also be useful to be able to compute

$$\int_0^{u_1} \cdots \int_0^{u_C} \dot{\mu}(z_1, \ldots, z_C, z_{C+1}, \ldots, z_D)\, \mathrm{d}z_1 \cdots \mathrm{d}z_C,$$

and a little thought shows that this equals (Problem 5.18)

$$-\frac{\partial^{D-C}}{\partial z_{C+1} \cdots \partial z_D} V(u_1, \ldots, u_C, z_{C+1}, \ldots, z_D). \tag{5.28}$$

Example 5.22 In Example 5.12, $D = 2$, and taking the function V found there we obtain $\partial^2 V/\partial z_1 \partial z_2 = -2/(z_1 + z_2)^3$, yielding

$$\dot{\mu}(z_1, z_2) = \frac{2}{(z_1 + z_2)^3}, \quad z_1, z_2 > 0, \quad \dot{\nu}(w) = -\frac{r^3}{2} \times \frac{-2}{r^3} = 1, \quad 0 < w < 1,$$

as expected. □

†This has to go later, or an example introducing this has to be earlier.

Example 5.23 (Hüsler–Reiss distribution) † Differentiation of (5.39) yields

$$-\frac{\partial^D V(z_1, \ldots, z_D)}{\partial z_1 \cdots \partial z_D} = \frac{1}{z_D^2 \prod_{d=1}^{D-1}(z_d \lambda_{d,D})} \phi_{D-1}(z^*_{-D}; \Lambda_{-D}),$$

where $\phi_{D-1}(\cdot; \Omega)$ is the $(D-1)$-variate Gaussian density with variance matrix Ω, and

$$z^*_{d;-D} = \frac{\lambda_{d,D}}{2} + \frac{\log z_d - \log z_D}{\lambda_{d,D}}, \quad d = 1, \ldots, D-1.$$

The angular density (5.27) therefore reduces to

$$\dot{\nu}(w) = \frac{1}{D w_D^2 \prod_{d=1}^{D-1}(w_d \lambda_{d,D})} \phi_{D-1}(w^*_{-D}; \Lambda_{-D}), \quad \sum_{d=1}^{D} w_d = 1, \quad w_d > 0,$$

where the dth element of w^*_{-D} is $\lambda_{d,D}/2 + \lambda^{-1}_{d,D} \log(w_d/w_D)$, and (5.28) is simply

$$\frac{1}{z_D^2 \prod_{d=C+1}^{D-1}(z_d \lambda_{d,D})} \frac{\partial^{D-C-1} \Phi_{D-1}(z^*_{-D}; \Lambda_{-D})}{\partial z_{C+1} \cdots \partial z_{D-1}},$$

with $z_1, \ldots, z_C$ replaced by $u_1, \ldots, u_C$. □

Construction of angular densities

The only constraint on the angular distribution ν is that every component of the angular variable W should have mean $1/D$, i.e., $\mathrm{E}(W) = D^{-1} 1_D$. In principle this gives a lot of liberty in the choice of ν, but it is not immediately obvious how to construct models that satisfy this constraint. One general approach to doing so starts from random variables $W^* = (W_1^*, \ldots, W_D^*)$ on the simplex $\mathcal{S}_{D-1}$ with probability density function $\dot{\nu}^*$ and arbitrary positive means

$$\mathrm{E}(W_d^*) = \int_{\mathcal{S}_{D-1}} w_d^* \dot{\nu}^*(w^*)\,\mathrm{d}w^* = m_d, \quad d = 1, \ldots, D.$$

As $\sum_d w_d^* = 1$, $\dot{\nu}^*$ depends on just $D-1$ of the w_d^*, and $\sum_d m_d = 1$. Now $\mathrm{E}(W_d^*/m_d) = 1$, suggesting that we represent the W_d^* in terms of new variables W_d that lie in $\mathcal{S}_{D-1}$, defined through the expressions

$$W_d = \frac{W_d^*/m_d}{\sum_{c=1}^D W_c^*/m_c}, \quad W_d^* = \frac{m_d W_d}{\sum_{c=1}^D m_c W_c}, \quad d = 1, \ldots, D. \tag{5.29}$$

The Jacobian $|\partial w^*/\partial w|$ equals $(m^{\mathrm{T}}w)^{-D} \prod_{d=1}^D m_d$, where $m^{\mathrm{T}}w = \sum_d w_d m_d$ is guaranteed to be positive. Thus $W = (W_1, \ldots, W_D)$ has probability density function

$$(m^{\mathrm{T}}w)^{-D} \left(\prod_{d=1}^D m_d \right) \dot{\nu}^* \left(\frac{m_1 w_1}{m^{\mathrm{T}}w}, \ldots, \frac{m_D w_D}{m^{\mathrm{T}}w} \right), \quad w \in \mathcal{S}_{D-1};$$

this depends on just $(D-1)$ of the w_d. For each d, $W_d^*/m_d = W_d/m^{\mathrm{T}}W$ has unit expectation, so

$$\begin{aligned} 1 &= \mathrm{E}\left(\frac{W_d}{m^{\mathrm{T}}W} \right) \\ &= \int_{\mathcal{S}_{D-1}} w_d (m^{\mathrm{T}}w)^{-(D+1)} \left(\prod_{d=1}^D m_d \right) \dot{\nu}^* \left(\frac{m_1 w_1}{m^{\mathrm{T}}w}, \ldots, \frac{m_D w_D}{m^{\mathrm{T}}w} \right) \mathrm{d}w, \end{aligned}$$

and on summing over d and recalling that $\sum_d w_d = 1$, we see that

$$\dot{\nu}(w) = D^{-1} (m^{\mathrm{T}}w)^{-(D+1)} \left(\prod_{d=1}^D m_d \right) \dot{\nu}^* \left(\frac{m_1 w_1}{m^{\mathrm{T}}w}, \ldots, \frac{m_D w_D}{m^{\mathrm{T}}w} \right)$$

is a probability density on $\mathcal{S}_{D-1}$ that satisfies

$$\int_{\mathcal{S}_{D-1}} w_d \dot{\nu}(w)\,\mathrm{d}w = D^{-1}, \quad d = 1, \ldots, D.$$

Thus a density $\dot{\nu}$ that satisfies the mean constraints is obtained by applying a change of variables to $\dot{\nu}^*$ and then tilting the result by dividing it by $Dm^{\mathrm{T}}w$. As this last factor is bounded, we can simulate from ν by generating W^* from ν^* and then accepting the transformed variables W given by (5.29) with probability $\min_d m_d/m^{\mathrm{T}}W$.

Example 5.24 (Dirichlet model) The Dirichlet density is

$$\dot{\nu}^*(w^*; \alpha) = \frac{\Gamma(\alpha.)}{\prod_d \Gamma(\alpha_d)} \prod_d (w_d^*)^{\alpha_d - 1}, \quad (w_1^*, \ldots, w_D^*) \in \mathcal{S}_{D-1},$$

where $\alpha_1, \ldots, \alpha_D > 0$ and $\alpha_{\cdot} = \sum_d \alpha_d$; here $m_d = \alpha_d/\alpha_{\cdot}$. The resulting tilted density,

$$\dot\nu(w;\alpha) = \frac{\Gamma(\alpha_{\cdot}+1)}{\left(\sum_d \alpha_d w_d\right)^{D+1}} \prod_{d=1}^{D} \left\{ \frac{\alpha_d}{\Gamma(\alpha_d)} \left(\frac{\alpha_d w_d}{\sum_c \alpha_c w_c} \right)^{\alpha_d - 1} \right\}, \quad w \in \mathcal{S}_{D-1},$$

is symmetric around its mean $D^{-1}1_D$ only when all the α_d are equal. The corresponding Pickands function is shown in Figure 5.3 for various values of α_1 and α_2.

This model can be extended using Dirichlet densities $\dot\nu_1^*, \ldots, \dot\nu_K^*$ with respective parameter vectors $\alpha^1, \ldots, \alpha^K$ and mean vectors $\alpha^k/\alpha_{\cdot}^k$. If we write $\dot\nu^*$ as a mixture of these densities with weights $p_1, \ldots, p_K$, i.e., W^* has density

$$\dot\nu^*(w^*;\alpha,p) = \sum_{k=1}^{K} p_k \dot\nu_k^*(w^*;\alpha^k), \quad p_1, \ldots, p_K > 0, \sum_{k=1}^{K} p_k = 1,$$

then in an obvious notation $\mathrm{E}(W_d^*) = m_d = \sum_k p_k \alpha_d^k/\alpha_{\cdot}^k$. This can be used to construct a tilted mixture that satisfies the mean constraint. □

Multivariate inference involves accounting for those features of the data that are key to the investigation at hand, and it is desirable that the key features can be directly related to particular model parameters. In many statistical settings these key features are first and second moments, but here the first moment is fixed, so dependence parameters are of primary interest.

- **Cooley et al. (2010)**
- **Ballini etc.**
- **inversion**

Angular distribution

The angular distribution can be a complex object, since ν could have densities and/or point masses on any combination of the $2^D - 1$ faces of $\mathcal{S}_{D-1}$. With $D = 3$, for example, there may be point masses on the corners $\{1,0,0\}$, $\{0,1,0\}$ and $\{0,0,1\}$, densities on the edges $\{(0,w_2,w_3) : w_2 + w_3 = 1\}$, $\{(w_1,0,w_3) : w_1 + w_3 = 1\}$ and $\{(w_1,w_2,0) : w_1 + w_2 = 1\}$, and on $S_2 = \{(w_1,w_2,w_3) : w_1 + w_2 + w_3 = 1\}$, in addition to point masses on S_2.† The only restrictions are that the total probability equal unity, and that $\mathrm{E}(W_1) = \mathrm{E}(W_2) = \mathrm{E}(W_3) = 1/3$.

†Make a figure for this.

A point mass p_1 at $\{1,0,0\}$ corresponds to extremes of Z_1 alone appearing with probability p_1, a total mass p_{23} on the face $\{(0,w_2,w_3) : w_2 + w_3 = 1\}$ corresponds to a probability p_{23} of large values of (Z_2,Z_3) occurring without large values of Z_1, with their relative sizes determined by the distribution on this face, and so forth. Since any of the faces of S_{D-1} is a lower-dimensional