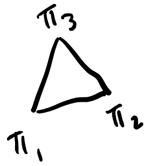
A Dirichlet distⁿ: this is a distⁿ of K-dimensional prob. simplex:

$$\Delta_K = \{(\pi_1, \pi_2, \ldots, \pi_K) : \pi_k \geq 0, \underbrace{\sum_k \pi_k = 1}\}.$$



$$\underline{\pi} = (\pi_1, \ldots, \pi_K) \sim \text{Dirichlet}(\alpha_1, \alpha_2, \ldots, \alpha_K)$$

assume $\alpha_1, \alpha_2, \ldots \alpha_K$ are known.

<u>Ex:</u> $\pi = (\pi_1, \pi_2) \sim \text{Dirichlet}(\alpha_1, \alpha_2)$

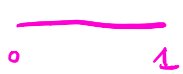
$$\text{Dirichlet}(\alpha_1,\alpha_2) = \text{Beta}(\alpha_1, \alpha_2)$$

$$p(\pi_1, \ldots, \pi_K) = \frac{\Gamma\left(\sum_k \alpha_k\right)}{\prod_k \Gamma(\alpha_k)} \prod_{k=1}^{K} \pi_k^{\alpha_k - 1}$$

$\pi_1 + \pi_2 = 1$

$$p(\pi_1, \pi_2) \propto \prod_{k=1}^{2} \pi_k^{\alpha_k - 1} = \pi_1^{\alpha_1 - 1} \pi_2^{\alpha_2 - 1}$$

$$= \underbrace{\pi_1^{\alpha_1 - 1} (1 - \pi_1)^{\alpha_2 - 1}}_{\text{Beta}(\alpha_1, \alpha_2)}$$



☆ inpt observation: The dirichlet is the multivariate version of the beta.

Let $\theta \sim \text{Dirchlet}(\overbrace{\alpha_1, \ldots, \alpha_K}^{\underset{\shortparallel}{\alpha}})$

then $p(\theta \mid \underline{\alpha}) \underset{\theta}{\propto} \prod_{K=1}^{m} \theta_k^{\alpha_k - 1}$,

$\sum_k \theta_K = 1, \quad \theta_K \geq 0 \quad \forall K.$

Good question: How do I set $\underline{\alpha}$?

$\underline{X} = (x_1, x_2, \ldots, x_n)$ where $x_i \in \{1, \ldots, m\}$.

Assume $\underline{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$; $\sum_i \theta_i = 1$.

$\underline{X} \mid \theta \overset{ind}{\sim} \text{Multinomial}(\underline{\theta})$

$\underline{X} \mid \theta \overset{ind}{\sim} \text{Categorical}(\underline{\theta})$

(0) $\boxed{P(X_i = j \mid \theta) = \theta_j = \theta_{x_i} \quad i=1,\ldots,n}$

This likelihood yields θ_j if $x_i = j$;

conjugate prior

o/w iid

Θ

$\hookrightarrow \theta_j^{I(x_i=j)}$

$$\underset{\sim}{\theta} \sim \text{Dir.}(\underline{\alpha}) \qquad \underline{\alpha} = (\alpha_1, \dots, \alpha_m)$$

$$\underbrace{p(\underline{\theta} \mid \underline{\alpha}) \underset{\theta}{\propto} \prod_{j=1}^{m} \theta_j^{\alpha_j - 1} \quad ; \quad \sum_j \theta_j = 1 ; \quad \theta_j \geq 0 \quad \forall j}_{\circledast}$$

Derive the posterior.

First, let's simplify $p(\underline{x} \mid \underline{\theta})$

$$p(\underline{x} \mid \theta) = \prod_{i=1}^{n} P(X_i = x_i \mid \theta)$$

by (0)

$$= \prod_{i=1}^{n} \theta_{x_i} = \theta_{x_1} \cdot \theta_{x_2} \cdots \theta_{x_n} \quad (1)$$

$$= \prod_{i=1}^{n} \prod_{j=1}^{m} \theta_j^{I(x_i=j)} \quad (2)$$

multin. likelih. (dep on i, j)

$$= \prod_{j=1}^{m} \left[\prod_{i=1}^{n} \theta_j^{I(x_i=j)} \right] \quad \text{(swapping products)} \quad (3)$$

$$= \prod_{j=1}^{m} \theta_j^{\sum_{i=1}^{n} I(x_i=j)} \quad \text{move } \prod_{i=1}^{n} \text{ into exponent} \quad (4)$$

Define $c_j = \# \{ i : x_i = j \}$

$$c = (c_1, \dots, c_m)$$

$$= \underbrace{\boxed{\prod_{j=1}^{m} \theta_j^{c_j}}}_{\text{will combine nicely with dirichlet prior.}} \qquad \text{by defn of } c_j \qquad (5)$$

$$p(\underline{x} \mid \underline{\theta}) \propto \prod_{i=1}^{m} \theta_j^{c_j}$$

$$p(\underline{\theta} \mid \underline{x}) \propto \overline{p(\underline{x} \mid \theta)}\, p(\underline{\theta})$$

$$\underset{\theta}{\propto} \prod_{j=1}^{m} \theta_j^{c_j} \times p(\theta) \qquad \text{by } (5)$$

$$\underset{\theta}{\propto} \prod_{j=1}^{m} \theta_j^{c_j} \times \prod_{j=1}^{m} \theta_j^{\alpha_j - 1}\, I\Big(\theta_j \geq 0, \sum_j \theta_j = 1\Big)$$

$$= \underbrace{\prod_{j=1}^{m} \theta_j^{\underline{c_j + \alpha_j} - 1}\, I\Big(\theta_j \geq 0, \sum_j \theta_j = 1\Big)}_{\text{Can you recognize this?}}$$

$$\underline{\theta} \mid \underline{x} \sim \mathrm{Dir}(\alpha_1 + c_1, \alpha_2 + c_2, \dots, \alpha_m + c_m)$$

$$= \mathrm{Dir}(\underline{\alpha} + \underline{c})$$