1 Definitions

- Symmetry $\sigma : \Theta \to \Theta$ is a measurable function with a measurable inverse satisfying

  $$\prod_k F_k(\theta) \propto \prod_k F_k\left(\sigma(\theta)\right)$$

  where the product is taken over non-prior terms. Symmetries must also fix the components of θ that correspond to observed variables.

- Local symmetry is a symmetry σ that satisfies

  $$F_k(\theta) \propto F_k\left(\sigma(\theta)\right)$$

  for all non-prior factors F_k.

1.1 Equivalence class

- A factor graph with equivalence relation Ξ, denoted $\mathcal{F} = (\theta, F, \Xi)$, induces a partition of Θ into equivalence classes $[\cdot]$, with the equivalence relation Ξ that there exists a symmetry between any two members of an equivalence class. Ξ specifies the type of symmetry.

  Denote $[\theta]$ to be the equivalence class of θ.

  Denote Θ/Ξ to be the set of all equivalence classes in Θ w.r.t Ξ, called the quotient set of Θ by Ξ.

  (TODO: local symmetry?)

- (TODO: Specific to breaking symmetry post-inference?)

  A symmetry detector $\Delta_{\mathcal{F}} : \Theta \to \Theta/\Xi$ is a measurable function that maps θ to $[\theta]$ w.r.t. Ξ.

- (TODO: Specific to breaking symmetry post-inference?)

  A symmetry breaker $\phi_{\mathcal{F}} : \Theta/\Xi \to \Theta$ is a measurable function that maps $[\theta]$ to θ_0 w.r.t. Ξ where θ_0 is a representative of $[\theta]$.

- (TODO: Specific to breaking symmetry post-inference?)

  An automatic symmetry breaker $\Phi_{\mathcal{F}} : \Theta \to \Theta$ is the composition $\Phi_{\mathcal{F}} = \phi_{\mathcal{F}} \circ \Delta_{\mathcal{F}}$ of a symmetry breaker with the corresponding symmetry detector.

  (TODO)

  An automatic symmetry breaker $\Phi_{\mathcal{F}} : \Theta \to \Theta$ is the composition $\Phi_{\mathcal{F}} = \phi_{\mathcal{F}} \circ \Delta_{\mathcal{F}}$ of a symmetry breaker with the corresponding symmetry detector that satisfies the following:

  1. For ASB for two types of symmetries Φ_1, Φ_2, $\Phi_1 \circ \Phi_2(\theta) = \Phi_2 \circ \Phi_1(\theta)$

  Has the following properties:

1. $\Phi \circ \Phi(\theta) = \Phi(\theta)$
2. $\forall \theta_* \in [\theta]$, $\Phi(\theta_*) = \Phi(\theta_0)$ where θ_0 representative of $[\theta]$

- A symmetry broken posterior distribution is a posterior distribution that has the (unnormalized) form

$$\prod_k F_k(\Phi_{\mathcal{F}_M} \circ ... \circ \Phi_{\mathcal{F}_1}(\theta))$$

where $\mathcal{F}_1, ..., \mathcal{F}_M$ have equivalence relations $\Xi_1, ..., \Xi_M$ corresponding to different types of symmetries equipped, respectively.

2 Breaking symmetries post-inference

Context: Have a posterior distribution $\prod_k F_k(\theta)$ where inference has been performed (θ has/can be identified)

Problem: Nonidentifiability. There exists a (local) symmetry σ s.t. $\prod_k F_k(\theta) \propto \prod_k F_k(\sigma(\theta))$ or for local, $F_k(\theta) \propto F_k(\sigma(\theta))$

When: Inference algorithm performance is not an issue in the presence of a symmetry.

Solution: Use the symmetry broken posterior distribution

2.1 Scaling symmetries

Scaling symmety is a symmetry σ which multiplies θ pointwise by a vector

$$v = (r_1, ..., r_N) = (e^{d_1}, ..., e^{d_N})$$

where $r_n \in \mathbb{R}_+$ and $d_n = \log r_n$.

- Let Σ_v be the diagonal matrix with v on the diagonal. Then $\sigma(\theta) = \Sigma_v \theta$.
- Let $d = (d_1, ..., d_N)$. Let $v_d = (e^{d_1}, ..., e^{d_N})$. Then the scaling symmetries of the model are $\sigma_d(\theta) = \Sigma_{v_d}\theta$, $d \in \mathcal{N}(C)$.
- Then $\{\Sigma_{v_d}\theta : d \in \mathcal{N}(C)\} \subset \Theta$ is an equivalence class for the model (θ, F).

3 Notes

- Class of transformations of model parameters $T \subset \{f : \Theta \to \Theta\}$. Subset of local symmetries contained in T $S_T \subset T$ defined by

$$S_T = \{\sigma \in T | \sigma \text{ is a local symmetry } \} = \bigcap_k \{\sigma \in T | F_k(\theta) \propto F_k(\sigma(\theta))\}$$

- Equivalence class definition

  A factor graph (model?) with equivalence relation Ξ, denoted $\mathcal{F} = (\theta, F, \Xi)$, induces a partition of Θ into equivalence classes, with the equivalence relation Ξ that there exists a symmetry between any two members of an equivalence class. Ξ specifies the type of symmetry.

  Classes may be singletons if the point is symmetric only to itself.

  Denote $[\theta]$ to be the equivalence class of θ.

  Denote Θ/Ξ to be the set of all equivalence classes in Θ w.r.t Ξ, called the quotient set of Θ by Ξ.

- A symmetry detector $\Delta_{\mathcal{F}} : \Theta \to \Theta/\Xi$ is a measurable? function that maps θ to $[\theta]$ w.r.t. Ξ.

  Symmetry detection is identifying the equivalence class?

- Definition 1: A symmetry breaker $\phi_{\mathcal{F}} : \Theta/\Xi \to \Theta$ is a measurable function that maps $[\theta]$ to θ_0 w.r.t. Ξ where θ_0 is a representative of $[\theta]$.

  Definition 2: A symmetry breaker $\phi_{\mathcal{F}} : \Theta/\Xi \to \Theta$ is a measurable function that maps $[\theta]$ to $\theta \in [\theta]$ w.r.t. Ξ.

- An automatic symmetry breaker $\Phi_{\mathcal{F}} : \Theta \to \Theta$ is the composition $\Phi_{\mathcal{F}} = \phi_{\mathcal{F}} \circ \Delta_{\mathcal{F}}$ of a symmetry breaker with the corresponding symmetry detector.

  A symmetry broken posterior distribution is a posterior distribution that has the (unnormalized) form

  $$\prod_k F_k(\Phi_{\mathcal{F}_M} \circ ... \circ \Phi_{\mathcal{F}_1}(\theta))$$

  where $\mathcal{F}_1, ..., \mathcal{F}_M$ have equivalence relations $\Xi_1, ..., \Xi_M$ corresponding to different types of symmetries equipped, respectively.

  Φ is not a symmetry (no measurable inverse)

- Scaling symmety is a symmetry σ which multiplies θ pointwise by a vector

  $$v = (r_1, ..., r_N) = (e^{d_1}, ..., e^{d_N})$$

  where $r_n \in \mathbb{R}_+$ and $d_n = \log r_n$.

  - Case of $r_n \in \mathbb{R}_-$ is covered by a combination of scaling symmetry and sign-flip symmetry.
  - Let Σ_v be the diagonal matrix with v on the diagonal. Then $\sigma(\theta) = \Sigma_v \theta$.
  - Let matrix C be the matrix of constraints on d_n. The scaling symmetries of the model are the vectors in the null space of C, $\mathcal{N}(C)$.
  - Let $d = (d_1, ..., d_N)$. Let $v_d = (e^{d_1}, ..., e^{d_N})$. Then the scaling symmetries of the model are $\sigma_d(\theta) = \Sigma_{v_d} \theta$, $d \in \mathcal{N}(C)$.

- Then $\{\Sigma_{v_d}\theta : d \in \mathcal{N}(C)\} \subset \Theta$ is an equivalence class for the model (θ, F).
- The symmetry detector $\Delta_{\mathcal{F}_S}$ for a scaling symmetry is defined as

$$\Delta_{\mathcal{F}_S}(\theta) = \{\Sigma_{v_d}\theta : d \in \mathcal{N}(C)\}$$

- A symmetry breaker (definition) for a scaling symmetry is

$$\phi_{\mathcal{F}_S}([\theta]) = \begin{cases} \theta & \mathcal{N}(C) = \{\vec{0}\} \\ \Sigma_{v_d}\theta,\ \|v_d\| = 1 & \mathcal{N}(C) \neq \{\vec{0}\} \end{cases}$$

- A symmetry breaker (definition 1) for a scaling symmetry is

$$\phi_{\mathcal{F}_S}([\theta]) = \Sigma_d\theta$$

  Let $\theta_* = \Sigma_{d^*}\theta$ for some $d^* \in \mathcal{N}(C)$.
  Then

$$\begin{aligned} \Delta_\Xi(\theta_*) &= \{\Sigma_d\theta_* : d \in \mathcal{N}(C)\} \\ &= \{\Sigma_d\Sigma_{d^*}\theta : d \in \mathcal{N}(C)\} \\ &= \{\Sigma_{d+d^*}\theta : d \in \mathcal{N}(C)\} \\ &= \{\Sigma_d\theta : d \in \mathcal{N}(C)\} \end{aligned}$$

  where the above follows as $d + d^* \in \mathcal{N}(C)$. Hence Δ_Ξ maps members of an equivalence class to the equivalence class.

  Symmetry breaker: need to constrain $d \in \mathcal{N}(C)$ to get a unique representative? How to remove dependence on θ?
- G&H 5.3: latent variable $z_i = X_i\beta + \epsilon_i$, $\epsilon_i \sim N(0, \sigma^2)$, scaling of β doesn't change sign of z_i. Proposed solution is just to fix σ.
- N'98: same problem as above but now covariance matrix. Proposed solution is just to fix $\Sigma_{1,1} = 1$.