MaxEnt model

Concept of Max Entropy Method

https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/how-presence-only-prediction-works.htm

Input data : presence only data (existing points and rasters)

Output: Relative occurrence rate: $P^*(z(x_i)) = \frac{exp(z(x_i)I)}{\sum exp(z(x_i)I)}$

where:

- z: vector of j variable at location x_i
- I: vector of regression coefficient
- x: pixel
- $z(x_i)I = z_1(x_i)^*\lambda_1 + z_2(x_i)^*\lambda_2 ... z_j(x_i)^*\lambda_j$

what we need to do is to find the best weight (λ) by using Max Entropy method, where the (information) entropy is calculate as:

Entropy: $H(p) = -\sum_{i=1}^{N} p(x_i) ln(p(x_i))$

where:

- N: Total pixels (in raster)
- $p(x_i)$: prob of species appearing at i pixel

Information entropy H(X) as shown in fig.1, actual value might way higher than 1 (this is an example for binary from wiki)

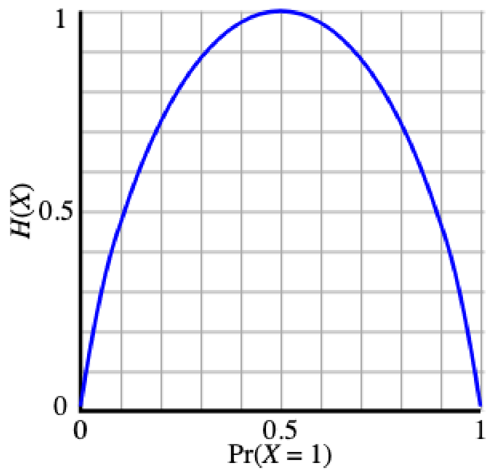

fig.1 Example curve

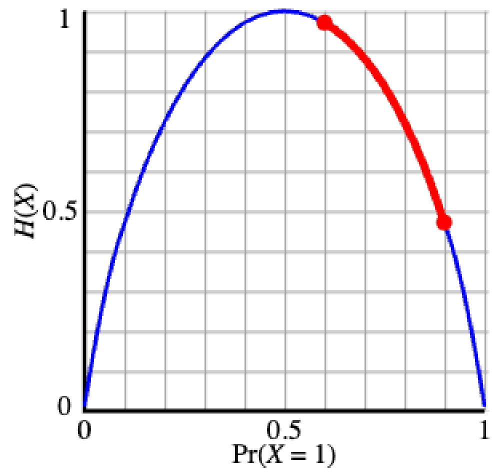

fig.2 What we want for constraint

As the entropy value H(X) increase, it interpret as the value is more evenly distributed, thus H(X) can be represent as fully guessing; If the value decrease to 0, it is too concentrate, or overfitting. So what we're finding is the max H(X) in a specific constraint as shown in figure 2.

→ In specific constraint we want the entropy to be more evenly distribute so that it is more reliable than making it lower H in that constraint.

Example for calculating Entropy for pixel

3 pixels (x_1, x_2, x_3), with 2 env variables (z_1, z_2)

where:

- z_1: water supplier, z_2: land value
- λ_1: 2, λ_2:-1 → The weight is I=[2,-1]
- Score: $z(x_i)I = z_1\lambda_1 + z_2\lambda_2$

pixel	(z_1, z_2)	product value $z(x_i)I$
x_1	[1, 5]	1 × 2 + 5 * (-1) = -3
x_2	[3, 3]	3 × 2 + 3 * (-1) = 3
x_3	[5, 1]	5 × 2 + 1 * (-1) = 9

After getting the score

$P^*(z(x_i)) = \frac{exp(score)}{\sum exp(score)}$ → $x_1 = e^{-3}, x_2 = e^3, x_3 = e^9$

⇒ $\sum exp(score) = e^{-3} + e^3 + e^9 \approx 8123$

⇒ $P(x_1) \approx 0, P(x_2) \approx 0.0148, P(x_3) \approx -0.0025$

⇒ H = - (0 - 0.0148 - 0.0025)=0.0173

→ Entropy is too small that we know it's too concentrate → change the weight (λ) to get a better value

> changing the weight is like make it fit the real world occurrence

Package

https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1541-0420.2012.01824.x?saml_referrer

Renner & Warton (2013) demonstrate that max entropy are mathematically equivalent to **generalised linear model (GLR) in poisson point process**

Similar to entropy value H, GLR is finding maximum likelihood by minimising deviance, which is also known as residual sum of square

> maximum likelihood: If the value from model is similar to actual value in the world

The R package maxnet is using this approximation for modelling MaxEnt by using glmnet, which includes different linear model