The Chain Rule for Mutual Information

Similarly to the **chain rule for entropy**, we can prove a chain rule for mutual information:

Corollary: Chain rule for mutual information

Let W, X, Y and Z be random variables. Then

$$I(WX;Y|Z) = I(X;Y|Z) + I(W;Y|ZX)\,.$$

Proof hint

Apply the **generalized chain rule**.

Show full proof

$$\begin{aligned} I(WX;Y|Z) &= H(WX|Z) - H(WX|YZ) \\ &= (H(X|Z) + H(W|XZ)) - (H(X|YZ) + H(W|XYZ)) \\ &= H(X|Z) - H(X|YZ) + H(W|XZ) - H(W|XYZ) \\ &= I(X;Y|Z) + I(W;Y|XZ). \end{aligned}$$