

Una rete feed-forward a livelli multipli

MLP (K_1, K_2)

In questa rete qui raffigurata dobbiamo poi inserire (in genere alla fine) la funzione di attivazione e la funzione softmax

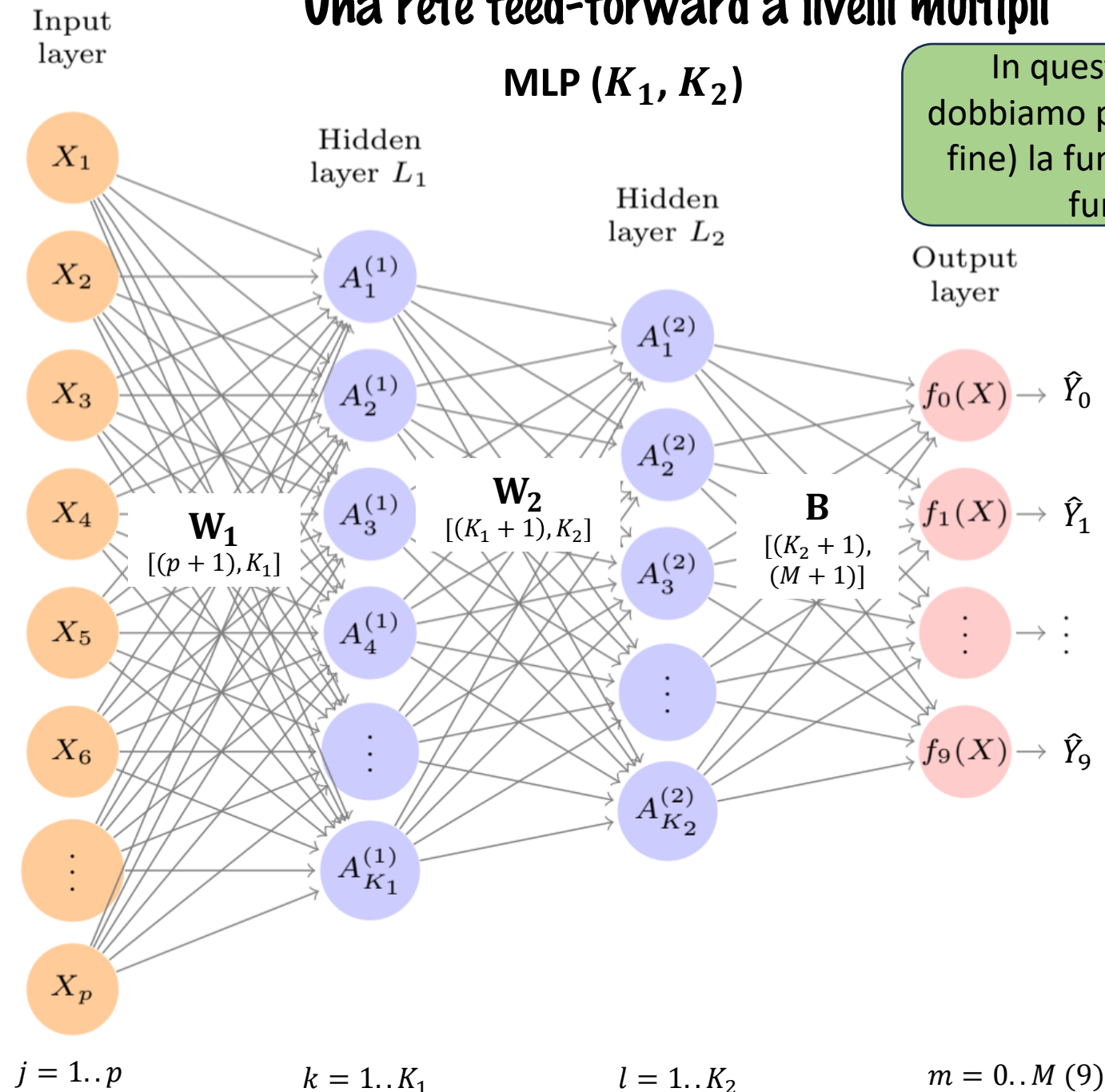
La matrice \mathbf{W}_1 contiene tutti i coefficienti che alimentano il primo livello nascosto L_1 dal livello di input.

La matrice \mathbf{W}_2 contiene tutti i coefficienti che alimentano il secondo livello nascosto L_2 dal primo livello nascosto L_1 .

La matrice \mathbf{B} contiene tutti i coefficienti che alimentano il livello di output dal secondo livello nascosto L_2 .

La prima dimensione di queste tre matrici è aumentata di 1 per tenere conto dei bias.

Non interpretare le notazioni \mathbf{W}_i e \mathbf{B} per weight e bias. Tutte le matrici contengono sia weight che bias.



le 10 variabili numeriche \hat{Y}_m rappresentano una singola variabile qualitativa

La matrice dei pesi \mathbf{W} di ogni layer è anche detta la «kernel matrix»