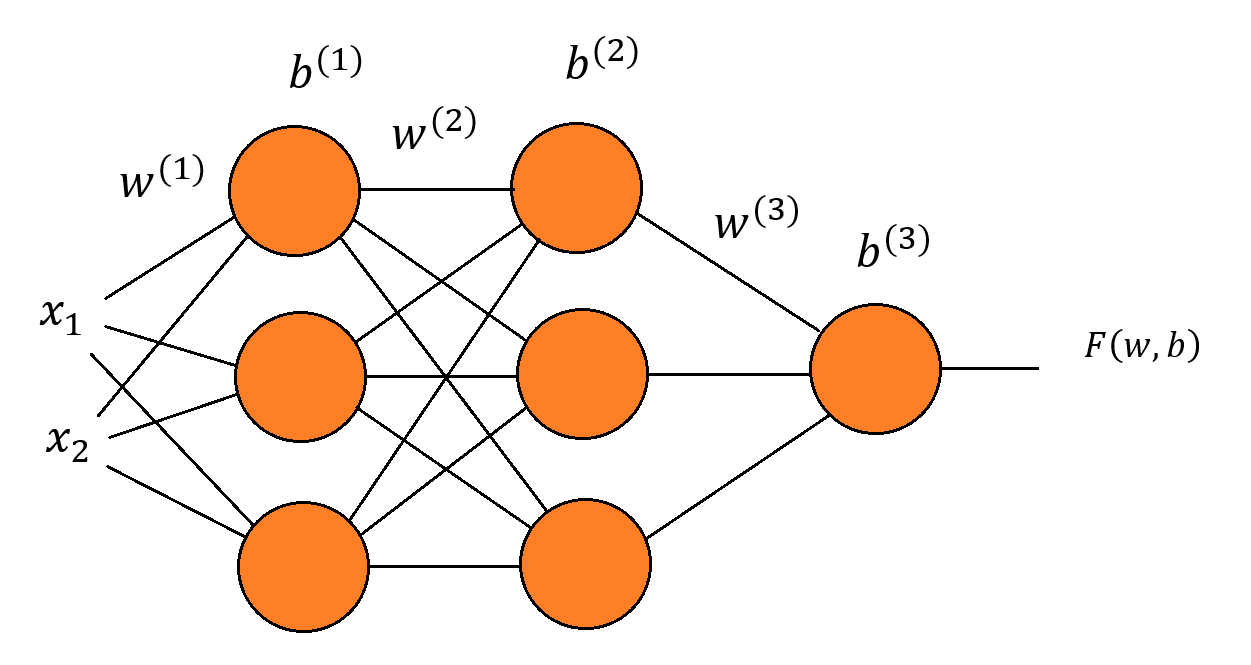
***Обратное распространение ошибки.***



Обозначим

– номер слоя

– параметры слоя

– входные данные слоя

– выходные данные слоя

Пусть L – число слоёв нейросети. Слой с номером L – выходной слой нейросети.

Целевая функция:

Результат обработки данных в скрытом слое:

– веса и смещения - го слоя, а выходные данные предыдущего слоя, индекс обозначает номер нейрона в слое.

– активационная функция. Ранее в качестве активационно мы использовали индикаторную функцию . Здесь и далее используем экспоненциальную сигмоиду:

Заметим, что .

Выход сети можно представить через результат работы предыдущего слоя

Активационная функция выходного слоя зависит от типа решаемой задачи. Для задачи классификации данных на два класса используем сигмоиду.

Для удобства изложения результат работы сети (как илюбого скрытого слоя) будем представлять в виде вектора. В рассматриваемом случае выходной вектор имеет одну компоненту .

Соединяя входы и выходы слоёв мы можем наращивать сложность нейросети. В рассматриваемом случае результат работы нейросети можно представить в следующей форме :

здесь *W* и *B* обозначают матрицу весов и вектор смещений соответствующего слоя сети.

***В этих и следующих формулах верхний индекс обозначает уровень нейросети, выходной уровень - последний.***

Обучение сети сводится к минимизации целевой функции или функции ошибок:

если *l*=0

при *l>*0

обозначим

если *l*=1

при *l>*1

если *l*=2

в матричной форме:

для произвольного значения *l*

Аналогичные рассуждения для производной по весу . Отличие при нахождении производной

Таким образом:

С учётом явного вида функции ошибок (потерь)

окончательно получаем:

или

Для производной по весу

где