

- 可以把层想象成由许多并行操作的单元组成，每个单元表示一个向量到标量的函数；

Q. 线性模型有缺点。

- A. • 该模型的表达能力局限在线性函数上，所以无法理解任何两个输入变量间的相互关系。

Q. 为扩展线性模型表示  $x$  的非线性函数

- A. • 可以把线性模型不用在  $x$  本身而是用在  $x$  变换后的输入  $\phi(x)$  上，这里  $\phi$  是非线性变换。

Q. 如何选择  $\phi$ 。

- A. •
1. 选择使用一个通用  $\phi$ ，例如无限维的  $\phi$ ，它隐含地用在基于 RBF 核
  2. 手动设计  $\phi$ 。
  3. 深度学习的自动学习  $\phi$ ， $y = f(x; \theta, w) = \phi(x; \theta)^T w$ 。
- 两种思路  $\left\{ \begin{array}{l} \text{用于从大量函数中学习 } \phi \text{ 的参考 } \theta, \\ \text{将 } \phi(x) \text{ 映射到需要的有限维的参考 } w \end{array} \right.$

Q. 深度学习的优点

- A. • 中  $\phi$  定义了一个隐藏层。这是三种方法中唯一一种放弃训练问题的凸性的方法，但是利大于弊。在这种方法中，我们将表示参数化为  $\phi(x; \theta)$ ，并且使用优化算法来寻找  $\theta$ ，使它能够得到一个好的表示。如果我们希望，这种方法也可以通过使它变得高度通用以获得第一种方法的优点——我们只需使用一个非常广泛的函数族  $\phi(x; \theta)$ 。这种方法也可以获得第二种方法的优点。人类专家可以将他们的知识编码进网络来帮助泛化，他们只需要设计那些他们期望能够表现优异的函数族  $\phi(x; \theta)$  即可。这种方法的优点是人类设计者只需要寻找正确的函数族即可，而不需要去寻找精确的函数。



