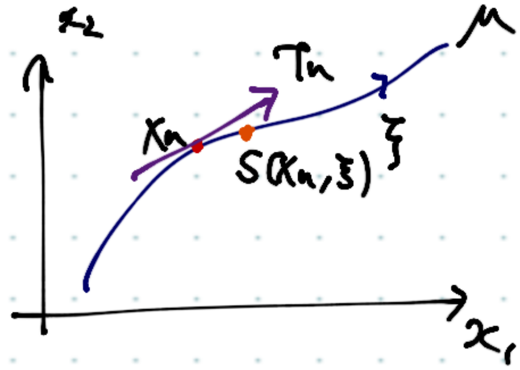
5.5.4 接線伝播法



$$\tau_n = \left.\frac{\partial s(x_n, \xi)}{\partial \xi}\right|_{\xi=0}$$

$$\left.\frac{\partial y_k}{\partial \xi}\right|_{\xi=0} = \sum_{i=1}^{D} \frac{\partial y_k}{\partial x_i} \left.\frac{\partial x_i}{\partial \xi}\right|_{\xi=0} = \sum_{i=1}^{D} J_{ki} \tau_i \quad (5.126)$$

接ベクトルを差分で近似

$$\tau_n = \frac{s(x_n, \xi) - s(x_n, 0)}{\xi}$$

正則化

$$\tilde{E} = E + \lambda \Omega \quad (5.127)$$

$$\Omega = \frac{1}{2} \sum_n \sum_k \left(\left.\frac{\partial y_{nk}}{\partial \xi}\right|_{\xi=0} \right)^2 = \frac{1}{2} \sum_n \sum_k \underline{\left(\sum_{i=1}^{D} J_{nki} \tau_{ni} \right)^2} \quad (5.128)$$

変換によるずれだけペナルティを課すことで出力を不変に保つようにさせる.

5.5.5 変換されたデータを用いた訓練

不変性を持たせる方法の1つ

— もとの入力パターンを変換して訓練集合を拡大する方法

と「接線伝播法」に関連があることを示す.

- 1つのパラメータ ξ による変換
- 二乗和誤差
- データ集合無限大

を考える.

- 1出力のNN

$$E = \frac{1}{2}\iint \{y(x) - t\}^2 \underbrace{p(t|x)p(x)}_{p(x|t)} dx\, dt \quad (1.87)$$

データを無限個コピーして分布 $\underline{p(\xi)}$ の ξ で摂動を受ける.

↑ 平均0で分散小さい ⇐ ずらしだけの変換

$$\tilde{E} = \frac{1}{2}\int E\, p(\xi)\, d\xi \quad (5.130)$$

$$s(x,\xi) = s(x,0) + \xi \frac{\partial}{\partial \xi} s(x,\xi)\Big|_{\xi=0} + \frac{\xi^2}{2}\frac{\partial^2}{\partial \xi^2} s(x,\xi)\Big|_{\xi=0} + O(\xi^3)$$

$$= x + \xi \tau + \frac{1}{2}\xi^2 \tau' + O(\xi^3)$$

変換された入力 $s(x,\xi)$ に対する出力

$$y(s(x,\xi)) = y(x) + (s(x,\xi) - x)^T \nabla y + \frac{1}{2}(s(x,\xi) - x)^T \nabla\nabla y\, (s(x,\xi) - x)$$
$$O((s(x,\xi) - x)^3)$$

$$= y(x) + \xi \tau^T \nabla y(x) + \frac{\xi^2}{2}\left[(\tau')^T \nabla y(x) + \tau^T \nabla\nabla y(x) \tau\right] + O(\xi^3)$$

$$[y(s(x,\xi)) - t]^2$$
$$= [y(x) - t]^2 + [y - t]\left[\xi \tau^T \nabla y + \frac{\xi^2}{2}\left[(\tau')^T \nabla y + \tau^T \nabla\nabla y\, \tau\right] + O(\xi^3)\right]$$
$$+ [\sim\sim]^2$$

$$\tilde{E} = \frac{1}{2}\iint \{y(x)-t\}^2 p(t|x)\,p(x)\,dx\,dt$$

$$+ \langle\xi\rangle \iint \{y(x)-t\}\,\tau^T \nabla y(x)\,p(t|x)\,p(x)\,dx\,dt$$

$$+ \langle\xi^2\rangle \frac{1}{2}\iint \Big[\{y-t\}\{(\tau')^T\nabla y + \tau^T \nabla\nabla y\,\tau\} + (\tau^T \nabla y)^2 \Big]$$

$$p(t|x)\,p(x)\,dx\,dt + O(\xi^3)$$

$\langle\xi\rangle = 0 \quad \langle\xi^2\rangle =: \lambda$

$$\Rightarrow \tilde{E} = E + \lambda\Omega \quad (5.131)$$

$$\Omega = \frac{1}{2}\int \Big[\{y - \mathbb{E}[t|x]\}\{(\tau')^T \nabla y(x) + \tau^T \nabla\nabla y(x)\,\tau\} + (\tau^T \nabla y)^2\Big] p(x)\,dx$$

$$y(x) = \mathbb{E}[t|x] + O(\xi^2)?$$

$\Rightarrow \Omega$の第1項は$O(\xi)$なので0

$$\Rightarrow \Omega = \frac{1}{2}\int (\tau^T \nabla y)^2 p(x)\,dx$$

接線伝播法の正則化(5.128)と同じ!!