



SCIENCE @ DIRECT®

Let  $\mathcal{X} \subset R^n$  and  $\mathcal{Y} \subset R^m$  be the input and output spaces, respectively. Let  $p(x, y) = p(x)p(y|x)$  and  $q(x, y) = q(x|y)q(y)$  be the joint and marginal distributions. Let  $\theta = (\theta_1, \dots, \theta_k) \in R^d$  be the parameters of the model. Let  $\alpha_j \geq 0$  and  $\sum_{j=1}^k \alpha_j = 1$  be the mixing coefficients. Let  $p_0(x) = (1/N) \sum_{t=1}^N \delta(x - x_t)$  be the empirical distribution. Let  $H(p||q)$  be the Kullback-Leibler divergence. Let  $J(\theta_k)$  be the cost function. Let  $\frac{\partial H(p||q)}{\partial \theta_k}$  be the gradient of the cost function. Let  $\frac{\partial H(p||q)}{\partial \theta_k} = \frac{1}{N} \sum_{t=1}^N \sum_{j=1}^k \frac{\alpha_j q(x_t|\theta_j)}{\sum_{i=1}^k \alpha_i q(x_t|\theta_i)} \alpha_j q(x_t|\theta_j)$ .

## 2. Gradient learning rule

Let  $x \in \mathcal{X} \subset R^n$  and  $y \in \mathcal{Y} \subset R^m$  be the input and output spaces, respectively. Let  $p(x, y) = p(x)p(y|x)$  and  $q(x, y) = q(x|y)q(y)$  be the joint and marginal distributions. Let  $\theta = (\theta_1, \dots, \theta_k) \in R^d$  be the parameters of the model. Let  $\alpha_j \geq 0$  and  $\sum_{j=1}^k \alpha_j = 1$  be the mixing coefficients. Let  $p_0(x) = (1/N) \sum_{t=1}^N \delta(x - x_t)$  be the empirical distribution. Let  $H(p||q)$  be the Kullback-Leibler divergence. Let  $J(\theta_k)$  be the cost function. Let  $\frac{\partial H(p||q)}{\partial \theta_k}$  be the gradient of the cost function. Let  $\frac{\partial H(p||q)}{\partial \theta_k} = \frac{1}{N} \sum_{t=1}^N \sum_{j=1}^k \frac{\alpha_j q(x_t|\theta_j)}{\sum_{i=1}^k \alpha_i q(x_t|\theta_i)} \alpha_j q(x_t|\theta_j)$ .

$$H(p||q) = \int p(y|x)p(x) \log \frac{p(y|x)p(x)}{q(x|y)q(y)} dx dy = z_q, \quad (1)$$

Let  $\theta = (\theta_1, \dots, \theta_k) \in R^d$  be the parameters of the model. Let  $\alpha_j \geq 0$  and  $\sum_{j=1}^k \alpha_j = 1$  be the mixing coefficients. Let  $p_0(x) = (1/N) \sum_{t=1}^N \delta(x - x_t)$  be the empirical distribution. Let  $H(p||q)$  be the Kullback-Leibler divergence. Let  $J(\theta_k)$  be the cost function. Let  $\frac{\partial H(p||q)}{\partial \theta_k}$  be the gradient of the cost function. Let  $\frac{\partial H(p||q)}{\partial \theta_k} = \frac{1}{N} \sum_{t=1}^N \sum_{j=1}^k \frac{\alpha_j q(x_t|\theta_j)}{\sum_{i=1}^k \alpha_i q(x_t|\theta_i)} \alpha_j q(x_t|\theta_j)$ .

$$p(y = j|x) = \frac{\alpha_j q(x|\theta_j)}{q(x|\Theta_k)}, \quad q(x|\Theta_k) = \sum_{j=1}^k \alpha_j q(x|\theta_j), \quad (2)$$

$$q(x|\theta_j) = q(x|y = j) = \theta_j \quad \text{g} \quad \text{z} \quad \text{a} \quad \text{z} \quad \text{a} \quad \text{z} \quad \text{a} \quad \Theta_k = \{\alpha_j, \theta_j\}_{j=1}^k. \quad (1), \quad \text{z} \quad \text{a}$$

$$H(p||q) = J(\Theta_k) = \frac{1}{N} \sum_{t=1}^N \sum_{j=1}^k \frac{\alpha_j q(x_t|\theta_j)}{\sum_{i=1}^k \alpha_i q(x_t|\theta_i)} \alpha_j q(x_t|\theta_j). \quad (3)$$

$$H(p||q) = \sum_{j=1}^k \int q(x|\theta_j) \ln \frac{q(x|\theta_j)}{p(x)} dx = \sum_{j=1}^k \int q(x|\theta_j) \ln q(x|\theta_j) dx - \sum_{j=1}^k \int q(x|\theta_j) \ln p(x) dx$$

$$q(x|\theta_j) = q(x|m_j, \Sigma_j) = \frac{1}{(\sqrt{2\pi})^n |\Sigma_j|^{1/2}} \exp\left(-\frac{1}{2}(x-m_j)^T \Sigma_j^{-1}(x-m_j)\right), \quad (1)$$

$$j = 1, 2, \dots, k, \quad -\infty < \beta_1, \dots, \beta_k < +\infty.$$

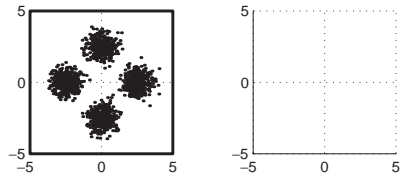
$\beta_j, m_j \in \mathbb{R}^n, \Sigma_j \in \mathbb{R}^{n \times n}$

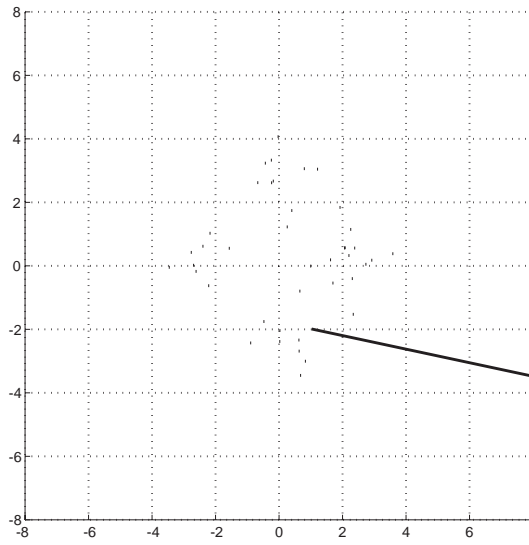
$$\Delta \beta_j = \eta \frac{\alpha_j}{N} \sum_{i=1}^k \sum_{t=1}^N h(i|x_t) U(i|x_t) (\delta_{ij} - \alpha_j), \quad (2)$$

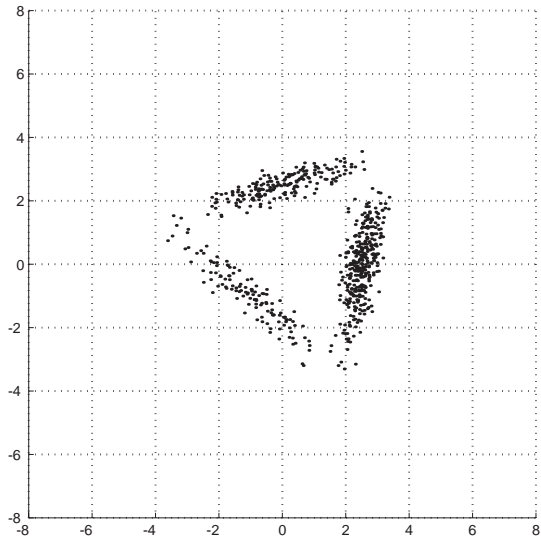
$$\Delta m_j = \eta \frac{\alpha_j}{N} \sum_{t=1}^N h(j|x_t) U(j|x_t) \Sigma_j^{-1} (x_t - m_j), \quad (3)$$

$$\Delta \Sigma_j = \eta \frac{\alpha_j}{2N} \sum_{t=1}^N h(j|x_t) U(j|x_t) \Sigma_j^{-1} (x_t - m_j)(x_t - m_j)^T - I \Sigma_j^{-1}, \quad (4)$$

$$U(i|x_t) = \sum_{r=1}^k (\delta_{ri} - p(r|x_t)) (\alpha_r q(x_t|\theta_r) + 1) x_t$$







$\theta_0$ ,  $\theta_1$ ,  $\theta_2$ ,  $\theta_3$ ,  $\theta_4$ ,  $\theta_5$ ,  $\theta_6$ ,  $\theta_7$ ,  $\theta_8$ ,  $\theta_9$ ,  $\theta_{10}$ ,  $\theta_{11}$ ,  $\theta_{12}$ ,  $\theta_{13}$ ,  $\theta_{14}$ ,  $\theta_{15}$ ,  $\theta_{16}$ ,  $\theta_{17}$ ,  $\theta_{18}$ ,  $\theta_{19}$ ,  $\theta_{20}$ ,  $\theta_{21}$ ,  $\theta_{22}$ ,  $\theta_{23}$ ,  $\theta_{24}$ ,  $\theta_{25}$ ,  $\theta_{26}$ ,  $\theta_{27}$ ,  $\theta_{28}$ ,  $\theta_{29}$ ,  $\theta_{30}$ ,  $\theta_{31}$ ,  $\theta_{32}$ ,  $\theta_{33}$ ,  $\theta_{34}$ ,  $\theta_{35}$ ,  $\theta_{36}$ ,  $\theta_{37}$ ,  $\theta_{38}$ ,  $\theta_{39}$ ,  $\theta_{40}$ ,  $\theta_{41}$ ,  $\theta_{42}$ ,  $\theta_{43}$ ,  $\theta_{44}$ ,  $\theta_{45}$ ,  $\theta_{46}$ ,  $\theta_{47}$ ,  $\theta_{48}$ ,  $\theta_{49}$ ,  $\theta_{50}$ ,  $\theta_{51}$ ,  $\theta_{52}$ ,  $\theta_{53}$ ,  $\theta_{54}$ ,  $\theta_{55}$ ,  $\theta_{56}$ ,  $\theta_{57}$ ,  $\theta_{58}$ ,  $\theta_{59}$ ,  $\theta_{60}$ ,  $\theta_{61}$ ,  $\theta_{62}$ ,  $\theta_{63}$ ,  $\theta_{64}$ ,  $\theta_{65}$ ,  $\theta_{66}$ ,  $\theta_{67}$ ,  $\theta_{68}$ ,  $\theta_{69}$ ,  $\theta_{70}$ ,  $\theta_{71}$ ,  $\theta_{72}$ ,  $\theta_{73}$ ,  $\theta_{74}$ ,  $\theta_{75}$ ,  $\theta_{76}$ ,  $\theta_{77}$ ,  $\theta_{78}$ ,  $\theta_{79}$ ,  $\theta_{80}$ ,  $\theta_{81}$ ,  $\theta_{82}$ ,  $\theta_{83}$ ,  $\theta_{84}$ ,  $\theta_{85}$ ,  $\theta_{86}$ ,  $\theta_{87}$ ,  $\theta_{88}$ ,  $\theta_{89}$ ,  $\theta_{90}$ ,  $\theta_{91}$ ,  $\theta_{92}$ ,  $\theta_{93}$ ,  $\theta_{94}$ ,  $\theta_{95}$ ,  $\theta_{96}$ ,  $\theta_{97}$ ,  $\theta_{98}$ ,  $\theta_{99}$ .

**4. Conclusions**

This paper presents a novel approach for... The proposed method... The results show... The proposed method... The results show...

**References**

1. J. Ma, ... (1) ...
2. ... (2) ...
3. ... (3) ...
4. ... (4) ...
5. ... (5) ...
6. ... (6) ...
7. ... (7) ...
8. ... (8) ...
9. ... (9) ...
10. ... (10) ...
11. ... (11) ...
12. ... (12) ...
13. ... (13) ...
14. ... (14) ...
15. ... (15) ...
16. ... (16) ...
17. ... (17) ...
18. ... (18) ...
19. ... (19) ...
20. ... (20) ...
21. ... (21) ...
22. ... (22) ...
23. ... (23) ...
24. ... (24) ...
25. ... (25) ...
26. ... (26) ...
27. ... (27) ...
28. ... (28) ...
29. ... (29) ...
30. ... (30) ...
31. ... (31) ...
32. ... (32) ...
33. ... (33) ...
34. ... (34) ...
35. ... (35) ...
36. ... (36) ...
37. ... (37) ...
38. ... (38) ...
39. ... (39) ...
40. ... (40) ...
41. ... (41) ...
42. ... (42) ...
43. ... (43) ...
44. ... (44) ...
45. ... (45) ...
46. ... (46) ...
47. ... (47) ...
48. ... (48) ...
49. ... (49) ...
50. ... (50) ...
51. ... (51) ...
52. ... (52) ...
53. ... (53) ...
54. ... (54) ...
55. ... (55) ...
56. ... (56) ...
57. ... (57) ...
58. ... (58) ...
59. ... (59) ...
60. ... (60) ...
61. ... (61) ...
62. ... (62) ...
63. ... (63) ...
64. ... (64) ...
65. ... (65) ...
66. ... (66) ...
67. ... (67) ...
68. ... (68) ...
69. ... (69) ...
70. ... (70) ...
71. ... (71) ...
72. ... (72) ...
73. ... (73) ...
74. ... (74) ...
75. ... (75) ...
76. ... (76) ...
77. ... (77) ...
78. ... (78) ...
79. ... (79) ...
80. ... (80) ...
81. ... (81) ...
82. ... (82) ...
83. ... (83) ...
84. ... (84) ...
85. ... (85) ...
86. ... (86) ...
87. ... (87) ...
88. ... (88) ...
89. ... (89) ...
90. ... (90) ...
91. ... (91) ...
92. ... (92) ...
93. ... (93) ...
94. ... (94) ...
95. ... (95) ...
96. ... (96) ...
97. ... (97) ...
98. ... (98) ...
99. ... (99) ...