

1. a.) Jun. 21. 2017 Q10.

b.) 對於 model 中一個 speaker 的所有參數(e.g.tri-phone, state, Gaussian, mean...)，把所有參數排成一個 vector。當 model 中有 L 個 speaker 的資料時，就可以生出 L 個 vector 的 data。

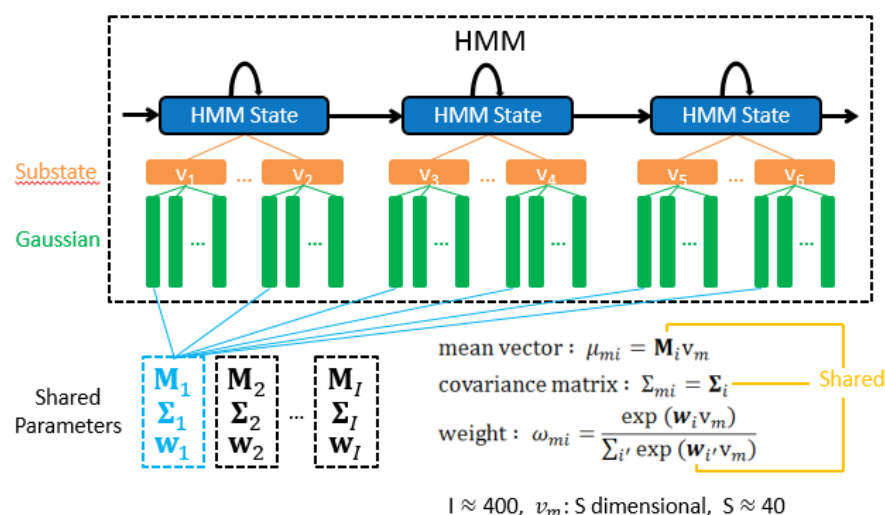
將每筆 data 減掉所有 data 的 mean，便會得到一組為 zero mean 的新 data，利用此筆 data 去做 PCA 運算降至 k 維後，就得到了 k 維的 eigenvoice space，同時也得到了 eigenvoice。

2. 在 CRF 中，假設每一個 word 都會有屬於它的詞類，而一個 word 的詞類會受到該 word 本身及前一個 word 的詞類所影響。對於 input 一個 sequence，可找出對應的詞類 sequence 的機率，求出機率最大的詞類 sequence 則可知道每一個 word 的詞類。

透過找出每一個 word 對應的詞類，並了解詞類的順序，就可以理解句子的文法，協助機器理解 user 在說什麼。

3. 對於一個 HMM state，把它拆成 N 個 sub-states，每個 substate 有 I 個 Gaussian 及自己 vector，令不同 substate 之間的 Gaussian 分享使用同一組 mean，covariance matrix，weight，所以會有 I 組 shared parameters。

對於不同 substate 之間的 gaussian，第 i 個 gaussian 的 mean 及 weight 是透過第 i 組 shared parameters 與該 substate 的 vector 運算求出來的。而第 i 個 gaussian 的 covariance matrix 都與第 i 組 shared parameters 的 covariance matrix 相同的。所以每個 substate 的每個 gaussian 都會是不一樣的。



4. 對於一篇文章，可以透過計算文章對應到不同 Topic 的機率，得出文章對應的 topic，同時可以計算一個 topic 對應到不同 word 的機率，得出一個 topic 應

該包含哪些 word。因此可以求出文章所對應的 Topic 及文章包含哪些 word。

5. VTS 的概念從 PMC 而來，當 model 在處理一個語音訊號時，同時去辨識出訊號中的 noise(e.g. 換氣時的雜音...)，馬上把找到的 noise 拿去調整 model，調整好的 model 就可以從該時間點開始去處理訊號。

假設 noise 在短時間內的變化不大，則每一個時間點的 model 都是用前一點的 noise 所調整出的 model。

對於 Clean Speech HMM's，把它從 Cepstral domain 逆運算回 Log-spectral domain，把 noise 看成是原來的訊號上一個小小的變化。利用 Taylor' Series，可以推算出對於 HMM 的參數加入一點變化後應該得到的值，利用此值作為包含 noise 情況的 HMM 的參數，然後運算回 Cepstral domain，就可得出調整完成的 Noisy speech HMM's 給 model 使用。

6. Jun. 21. 2017 Q5.

7. 了解 Spoken document 的內容，並從中取得有用的資訊，例如從 document 中找出 key words/key term/topic...，從而協助系統完成不同工作，例如 generate Summarization/Title，利用 topic/key term 來為 document 進行分類....。

8. 在 binary classifier 問題上，SVM 可以找出從兩個 class 的 data 中找出一個最佳 hyperplane separating(分界線/分界面)，令兩個 class 的點與分界線/分界面有最遠的距離，即兩個 class 可以分得最開。

對於一些非線性可分的 data，可以透過 Feature Mapping 的方法提高維度，令 data 在高維度中線性可分。

對於包含 noise data 的情況，可以 soft margin 的方式，不考慮一些跟 class 裡整體 data 分佈情況不像的 data，達到把 noise data 排除的效果。

9. Jun. 21. 2017 Q6.

10. Mismatch: train data 及 test data 並非一致，很多情況下 train data 是沒有 noise，而 test data 是有 noise 的，因此做出來的 model 效果會不好。(noise: OOV/語氣/情緒/背景音...)

Model-based: 從 model 入手解決 noise 問題，假設 noise 在短時間內的變化不大，接 Q5 第一段。Example: PMC, VTS

Feature-based: 從 Feature 入手解決 noise 問題，假如 noise 在 time domain 上與 data 做 convolutional，在 time domain 上就會是 data 加上 noise，若我們可以知道 noise 在 time domain 上的值，則可以直接在 time domain 上減掉 noise，從而解決 noise 的問題。Example: CMS, CMVN, HEQ