Solutions to CS6740/INFO6300 prior final exam

1a

Note that $\vec{w} = \begin{pmatrix}0\\1\end{pmatrix}$ means: ignore $\cos(\vec{q},\vec{d})$, just prefer docs w/ a higher class #. (+1)

(not mentioning may mean ignore)

(~~[illegible]~~)

The Joachims' alg. creates a pref. constraint for any pair of docs (summaries) s_i, s_j s.t. s_i is higher on the pg, s_i is not clicked on, and s_j is. (explanation)

$\Rightarrow$ we ~~only~~ infer that d_2^b is preferred over d_1^b. (+1)

(+1) inference: but, due to presentation bias, we do not infer that d_1^a is preferred over d_2^a despite the click pattern. (+1)

omitted coords for $v^{(4)} \ldots v^{(6)}$ since they're not relevant to the q_b situation (all coords for them $= 0$)

now, $\vec{d}_1^b$ = norm'd version of

$$\begin{bmatrix}1\\0\\1\\1\\0\\0\end{bmatrix} = \cancel{\begin{bmatrix}1/\sqrt{3}\\0\\1/\sqrt{3}\\1/\sqrt{3}\\0\\0\end{bmatrix}} \rightarrow \begin{bmatrix}1/\sqrt{3}\\0\\1/\sqrt{3}\\1/\sqrt{3}\\0\\0\end{bmatrix}$$

(+1): term counts

(+1) normalization. (or length captation [illegible])

$\vec{d}_2^b$ = norm'd version of

$$\begin{bmatrix}0\\0\\1\\0\\1\\1\end{bmatrix} = \begin{bmatrix}0\\0\\1/\sqrt{3}\\0\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$$

$$\vec{q}_b: \begin{bmatrix}1\\1\\1\\0\\0\\0\end{bmatrix} \rightarrow \text{norm'd} = \begin{bmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\\0\\0\\0\end{bmatrix}$$

$$\Rightarrow \cos(\vec{q}_b, \vec{d}_1^b) = \frac{1}{3} + \frac{1}{3} = 2/3$$

$$\cos(\vec{q}_b, \vec{d}_2^b) = 1/3.$$

(+1): compute cosine

$$\Rightarrow \phi(\vec{q}_b, \vec{d}_1^b) = \begin{bmatrix}2/3\\-1\end{bmatrix}, \quad \phi(\vec{q}_b, \vec{d}_2^b) = \begin{bmatrix}1/3\\1\end{bmatrix}$$

since dotting w/ $\vec{w} = \begin{pmatrix}0\\1\end{pmatrix}$ simply yields the second coordinate, $\vec{w} = \begin{pmatrix}0\\1\end{pmatrix}$ does indeed rank $\vec{d}_2^b$ over $\vec{d}_1^b$ wrt $\vec{q}_b$

(+1): what does ~~check~~ $\vec{w}$ do ~~[illegible]~~ (~~[illegible]~~ ~~[illegible]~~ ~~can still get this pt~~) in terms of ranking the docs wrt the queries.

= 7 points possible

(only subtract pts for egregious math mistakes – want to stress concepts)

1 b.

w/ non-normalized wt vectors: computed in previous response.

$$\vec{q}^{\,new} = \alpha \cdot \vec{q}_b + \frac{\beta}{\# \cancel{rel}\ \vec{d}_i\ rel} \cdot \sum_{\vec{d}_i\ rel} \vec{d}_i + \underset{\substack{\shortparallel \\ 0}}{\gamma} \ldots$$

0 - ignoring neg feedback.

$$= \vec{q}_b + \beta \underset{\uparrow \text{ un'norm.}}{\vec{d}_2^{\,b}}$$

(+1) (understand Rocchio eqn)

$$= \begin{bmatrix} 1 \\ 1 \\ 1+\beta \\ 0 \\ \beta \\ \beta \end{bmatrix}$$

(+1) compute ~~the new q~~ $\vec{q}^{\,new}$

for un'norm'd versions of vectors:

$$\underbrace{\vec{q}^{\,new} \cdot \vec{d}_1^{\,b} = \begin{bmatrix} 1 \\ 1 \\ 1+\beta \\ 0 \\ \beta \\ \beta \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = 1 + 1 + \beta \qquad \vec{q}^{\,new} \cdot \vec{d}_2^{\,b} = \begin{bmatrix} 1 \\ 1 \\ 1+\beta \\ 0 \\ \beta \\ \beta \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}}_{} = 1+\beta+\beta+\beta = 1+3\beta$$

(+1) (recomputing) correct doc vectors)

find β s.t. $2+\beta < 1+3\beta$

$\equiv \quad 1 < 2\beta$

$\equiv \quad \frac{1}{2} < \beta$

(+1) solve for β (ok if slightly off.)

= 4 pts total.

2 :

a $P(R=y \mid D=d) = \sum_{t_i} P(T=t_i, R=y \mid D=d)$ (introduce r.v.) (+1) ~~Bayes flip~~ introduce r.v.

$= P(T=t_1 \mid D=d) + P(T=t_2 \mid D=d)$ (def. of relevance) (+1) def. of relevance

$= \frac{P(D=d \mid T=t_1) P(T=t_1)}{P(D=d)} + \frac{P(D=d \mid T=t_2) P(T=t_2)}{P(D=d)}$ (+1) Bayes flip.

can assume unif. (+1) [problem subsequently rewritten to make this assumption given]

from given topic model (+1)

$= \sum_t P(T=t, D=d)$ is one possibility (+1). Or ~~assume unif.~~ (or, assume unif.: ok since 'sol'ns' say r.v.) ~~no credit for unif.~~ ~~(they're seen this before)~~

6 pts total.

~~b $\sum_t P(T=t, R=y \mid D=d)$~~ ~~not really an appropriate r.v.~~

3 a. SVD: $U \Sigma V^T$.

Since U orthonormal,

$\begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ a \\ b \end{bmatrix} = 0 \Rightarrow a = 0$. (+1)

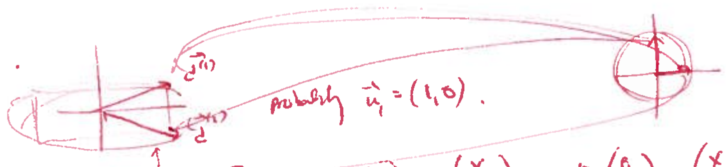
and length of $\vec{u}_2 = 1 \Rightarrow b = \pm 1$. (+1)

$\boxed{6}$ by defn, Σ is diagonal $\Rightarrow c = d = 0$ (+1) (ok if strictly greater)

also, by convention, $0 < e \leq 5$ (+1)(+1)

finally, V must have 2 columns b/c U does, but we don't know how many rows it has. (+1) [(+)0.5 if says we know the "m"].

b.



probably $\vec{u}_1 = (1, 0)$.

$D = \begin{bmatrix} x & x \\ y & -y \end{bmatrix}$ $D\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ $D\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$. lengths $= \sqrt{x^2+y^2}$ (+1) <told>

split diff: must be $\begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix}$ (+1) (length norm).

$D\begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}x \\ \frac{\sqrt{2}}{2}y - \frac{\sqrt{2}}{2}y \end{bmatrix} = \begin{bmatrix} \sqrt{2}x \\ 0 \end{bmatrix}$ length $= \sqrt{2}x = \sqrt{2x^2} > \sqrt{x^2+y^2}$.

$\boxed{5}$

(+1) <what $\vec{u}_1$ corresponds to> (+1) (length comparison).

so, since ~~left singular~~ $\vec{u}_1$ corresponds to largest axis of semicircle; we know ~~this vector~~ its length must be $\geq \sqrt{2}x$, since ... this direct... (+1) right conclusion

4. Lagrange multiplier eqn:

$$\sum_t P_{\theta_{i-1}}(t|w)\, K(t) \prod_j \rho_j^{\#(\text{rule } j \text{ in } t)} - \underbrace{\lambda \sum_j (\rho_j - 1)}_{+1 \text{ L.m.}}$$

(+.5 if eg. $\Sigma K(t) \rho_j$

pick specific ρ_1 (for simplicity), take derivative $\frac{\partial}{\partial \rho_1}$:

$$\sum_t P_{\theta_{i-1}}(t|w)\, K(t) \boxed{\#(\text{rule } 1 \text{ in } t)\, \rho_1^{\#(\text{rule } 1 \text{ in } t)-1}} \prod_{j \neq 1} \rho_j^{\#(\text{rule } j \text{ in } t)} - \lambda .$$

constant wrt ρ_1

+1 this should be in here.

+1 ~~derivative~~ $P_{\theta_{i-1}}$ not a fn of ρ_1.

+1

~~problem~~

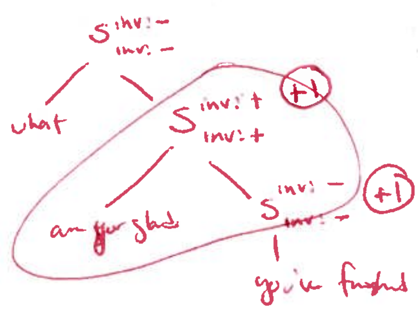
problem: can't ~~solve for ρ_1~~ solve for ρ_1: has exponent depending on t (so can't take out of sum) and ρ_1 soln depends on the other ρ_j's.

+1 +1

6 total

5. (a) adjoin . β_{says} into α_{wff} @ node 2: +1 (adj into correct node)

6 total



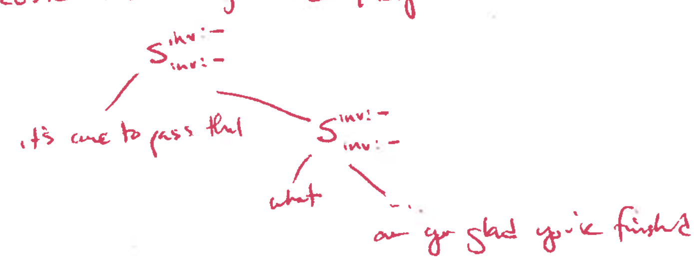
~~no showing features @ any node so derivation is complete.~~

(feature propagated right)

(b) tree α_{wff} has incompatible feature at node 2. +1

(c). Kind that node 0 is suspicious, also haven't used β_{itgot}:

could ~~insert~~ adjoin: β_{itg} into node 0 of the above derived tree:



+1 finish correct.s,

+1 ~~explain what problem is~~ (allowed top node in further adjunction)

~~contradiction~~ ~~propagate~~ what the problem is. ~~all nodes have perfectly fine or strike fail top/bottom features.~~

(e.g. inv:− inv:− ~~so so this should be fine~~