



# Variational Inference

Material adapted from David Blei  
University of Maryland

EXERCISE

## Feedback

- Disagreement from you: flipped classroom (but most like), theory
- I agree with you: one size fits all, inconsistent notation, reading inconsistent, questions in random order
- Sorry: Slides w/ video, hard to hear questions, too much NLP
- Huh?: Online judge, ask questions on same interface, A cutoff, video speed
- Thinking about: second midterm

## Announcements

- Next class: unflipped VAE / GAN
- Reading: Goodfellow
- Next week: LSTM / RNN
- Rest of course remains unchanged
- Default project: more data, Kaggle site this week

## Useful Reference: $\Psi(x)$

$x$	$\Psi(x)$
1	-0.577215664902
2	0.422784335098
3	0.922784335098
4	1.25611766843
5	1.50611766843
6	1.70611766843
7	1.8727843351
8	2.01564147796
9	2.14064147796

Also: from `scipy.special` import `digamma`

## Example

- Three topics, same documents as last time

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix} \quad (1)$$

- Assume uniform  $\gamma$ :  $(2.0, 2.0, 2.0)$
- Compute update for  $\phi$

$$\phi_{ni} \propto \beta_{iv} \exp \left( \Psi(\gamma_i) - \Psi \left( \sum_j \gamma_j \right) \right) \quad (2)$$

- For a the first word (dog) in the document: **dog cat cat pig**

## Example

- Three topics, same documents as last time

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix} \quad (1)$$

- Assume uniform  $\gamma$ :  $(2.0, 2.0, 2.0)$
- Compute update for  $\phi$

$$\phi_{ni} \propto \beta_{iv} \exp \left( \Psi(\gamma_i) - \Psi \left( \sum_j \gamma_j \right) \right) \quad (2)$$

- For a the first word (dog) in the document: **dog cat cat pig**

## Update $\phi$ for dog

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix}$$

$$\phi_{ni} \propto \beta_{iv} \exp \left( \Psi(\gamma_i) - \Psi \left( \sum_j \gamma_j \right) \right)$$

- $\gamma = (2.000, 2.000, 2.000)$

## Update $\phi$ for dog

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix} \quad \begin{aligned} \phi_{ni} &\propto \\ \beta_{iv} &\exp\left(\Psi(\gamma_i) - \Psi\left(\sum_j \gamma_j\right)\right) \end{aligned}$$

- $\gamma = (2.000, 2.000, 2.000)$
- $\phi(0) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$



## Update $\phi$ for dog

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix} \quad \begin{aligned} \phi_{ni} &\propto \\ \beta_{iv} \exp\left(\Psi(\gamma_i) - \Psi\left(\sum_j \gamma_j\right)\right) \end{aligned}$$

- $\gamma = (2.000, 2.000, 2.000)$
- $\phi(0) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- $\phi(1) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$

## Update $\phi$ for dog

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix} \quad \begin{aligned} \phi_{ni} &\propto \\ \beta_{iv} &\exp\left(\Psi(\gamma_i) - \Psi\left(\sum_j \gamma_j\right)\right) \end{aligned}$$

- $\gamma = (2.000, 2.000, 2.000)$
- $\phi(0) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- $\phi(1) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- $\phi(2) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$

## Update $\phi$ for dog

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix} \quad \begin{aligned} \phi_{ni} &\propto \\ \beta_{iv} \exp\left(\Psi(\gamma_i) - \Psi\left(\sum_j \gamma_j\right)\right) \end{aligned}$$

- $\gamma = (2.000, 2.000, 2.000)$
- $\phi(0) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- $\phi(1) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- $\phi(2) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- After normalization:  $\{0.333, 0.333, 0.333\}$

## Update $\phi$ for pig

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix}$$

$$\phi_{ni} \propto$$

$$\beta_{iv} \exp\left(\Psi(\gamma_i) - \Psi\left(\sum_j \gamma_j\right)\right)$$

- $\gamma = (2.000, 2.000, 2.000)$

## Update $\phi$ for pig

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix} \quad \begin{aligned} \phi_{ni} &\propto \\ \beta_{iv} \exp\left(\Psi(\gamma_i) - \Psi\left(\sum_j \gamma_j\right)\right) \end{aligned}$$

- $\gamma = (2.000, 2.000, 2.000)$
- $\phi(0) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$

## Update $\phi$ for pig

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix} \quad \begin{aligned} \phi_{ni} &\propto \\ \beta_{iv} &\exp\left(\Psi(\gamma_i) - \Psi\left(\sum_j \gamma_j\right)\right) \end{aligned}$$

- $\gamma = (2.000, 2.000, 2.000)$
- $\phi(0) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- $\phi(1) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$

## Update $\phi$ for pig

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix} \quad \begin{aligned} \phi_{ni} &\propto \\ \beta_{iv} &\exp\left(\Psi(\gamma_i) - \Psi\left(\sum_j \gamma_j\right)\right) \end{aligned}$$

- $\gamma = (2.000, 2.000, 2.000)$
- $\phi(0) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- $\phi(1) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- $\phi(2) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$

## Update $\phi$ for pig

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix} \quad \begin{aligned} \phi_{ni} &\propto \\ \beta_{iv} \exp\left(\Psi(\gamma_i) - \Psi\left(\sum_j \gamma_j\right)\right) \end{aligned}$$

- $\gamma = (2.000, 2.000, 2.000)$
- $\phi(0) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- $\phi(1) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- $\phi(2) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- After normalization:  $\{0.333, 0.333, 0.333\}$



## Update $\phi$ for cat

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix}$$

$$\phi_{ni} \propto$$

$$\beta_{iv} \exp\left(\Psi(\gamma_i) - \Psi\left(\sum_j \gamma_j\right)\right)$$

- $\gamma = (2.000, 2.000, 2.000)$

## Update $\phi$ for cat

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix} \quad \begin{aligned} \phi_{ni} &\propto \\ \beta_{iv} \exp\left(\Psi(\gamma_i) - \Psi\left(\sum_j \gamma_j\right)\right) \end{aligned}$$

- $\gamma = (2.000, 2.000, 2.000)$
- $\phi(0) \propto 0.260 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.072$

## Update $\phi$ for cat

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix} \quad \begin{aligned} \phi_{ni} &\propto \\ \beta_{iv} \exp\left(\Psi(\gamma_i) - \Psi\left(\sum_j \gamma_j\right)\right) \end{aligned}$$

- $\gamma = (2.000, 2.000, 2.000)$
- $\phi(0) \propto 0.260 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.072$
- $\phi(1) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$

## Update $\phi$ for cat

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix} \quad \begin{aligned} \phi_{ni} &\propto \\ \beta_{iv} &\exp\left(\Psi(\gamma_i) - \Psi\left(\sum_j \gamma_j\right)\right) \end{aligned}$$

- $\gamma = (2.000, 2.000, 2.000)$
- $\phi(0) \propto 0.260 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.072$
- $\phi(1) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- $\phi(2) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$

## Update $\phi$ for cat

$$\beta = \begin{bmatrix} \text{cat} & \text{dog} & \text{hamburger} & \text{iron} & \text{pig} \\ .26 & .185 & .185 & .185 & .185 \\ .185 & .185 & .26 & .185 & .185 \\ .185 & .185 & .185 & .26 & .185 \end{bmatrix} \quad \begin{aligned} \phi_{ni} &\propto \\ \beta_{iv} \exp\left(\Psi(\gamma_i) - \Psi\left(\sum_j \gamma_j\right)\right) \end{aligned}$$

- $\gamma = (2.000, 2.000, 2.000)$
- $\phi(0) \propto 0.260 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.072$
- $\phi(1) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- $\phi(2) \propto 0.185 \times \exp(\Psi(2.000) - \Psi(2.000 + 2.000 + 2.000)) = 0.051$
- After normalization:  $\{0.413, 0.294, 0.294\}$

## Update $\gamma$

- Document: dog cat cat pig
- Update equation

$$\gamma_i = \alpha_i + \sum_n \phi_{ni} \quad (3)$$

- Assume  $\alpha = (.1, .1, .1)$

## Update $\gamma$

- Document: dog cat cat pig
- Update equation

$$\gamma_i = \alpha_i + \sum_n \phi_{ni} \quad (3)$$

- Assume  $\alpha = (.1, .1, .1)$

	$\phi_0$	$\phi_1$	$\phi_2$
dog	.333	.333	.333
cat	.413	.294	.294
pig	.333	.333	.333
$\alpha$	0.1	0.1	0.1
sum	1.592	1.354	1.354

- Note: **do not normalize!**

## Update $\gamma$

- Document: dog cat cat pig
- Update equation

$$\gamma_i = \alpha_i + \sum_n \phi_{ni} \quad (3)$$

- Assume  $\alpha = (.1, .1, .1)$

	$\phi_0$	$\phi_1$	$\phi_2$	
dog	.333	.333	.333	
cat	.413	.294	.294	x2
pig	.333	.333	.333	
$\alpha$	0.1	0.1	0.1	
sum	1.592	1.354	1.354	

- Note: **do not normalize!**



## Update $\beta$

- Count up all of the  $\phi$  across all documents
- For each topic, divide by total
- Corresponds to maximum likelihood of expected counts

## Update $\beta$

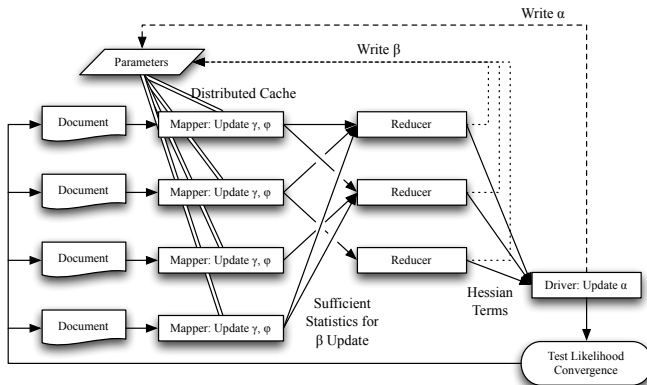
- Count up all of the  $\phi$  across all documents
- For each topic, divide by total
- Corresponds to maximum likelihood of expected counts
- Unlike Gibbs sampling, no Dirichlet prior

## Automatic Inference



## Parallel LDA

Zhai et al, 2012



## Online LDA

Hoffman and Blei, 2010

---

**Algorithm 2** Online variational Bayes for LDA

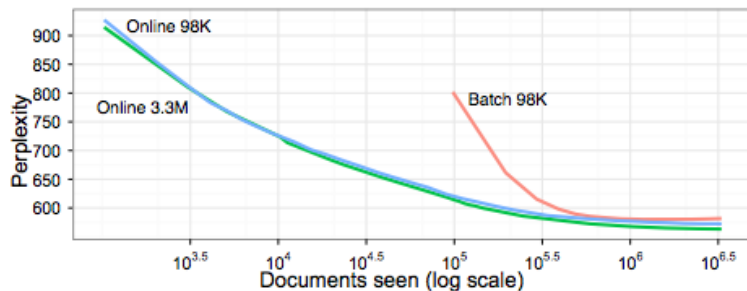
---

Define  $\rho_t \triangleq (\tau_0 + t)^{-\kappa}$   
 Initialize  $\lambda$  randomly.  
**for**  $t = 0$  to  $\infty$  **do**  
   *E step:*  
   Initialize  $\gamma_{tk} = 1$ . (The constant 1 is arbitrary.)  
   **repeat**  
     Set  $\phi_{twk} \propto \exp\{\mathbb{E}_q[\log \theta_{tk}] + \mathbb{E}_q[\log \beta_{kw}]\}$   
     Set  $\gamma_{tk} = \alpha + \sum_w \phi_{twk} n_{tw}$   
   **until**  $\frac{1}{K} \sum_k |\text{change in } \gamma_{tk}| < 0.00001$   
   *M step:*  
   Compute  $\tilde{\lambda}_{kw} = \eta + D n_{tw} \phi_{twk}$   
   Set  $\lambda = (1 - \rho_t) \lambda + \rho_t \tilde{\lambda}$ .  
**end for**

---

## Online LDA

Hoffman and Blei, 2010



## Online LDA

Hoffman and Blei, 2010

- 1: Initialize  $\lambda^{(0)}$  randomly.
- 2: Set the step-size schedule  $\rho_t$  appropriately.
- 3: **repeat**
- 4:   Sample a data point  $x_i$  uniformly from the data set.
- 5:   Compute its local variational parameter,

$$\phi = \mathbb{E}_{\lambda^{(t-1)}}[\eta_g(x_i^{(N)}, z_i^{(N)})].$$

- 6:   Compute intermediate global parameters as though  $x_i$  is replicated  $N$  times,

$$\hat{\lambda} = \mathbb{E}_{\phi}[\eta_g(x_i^{(N)}, z_i^{(N)})].$$

- 7:   Update the current estimate of the global variational parameters,

$$\lambda^{(t)} = (1 - \rho_t)\lambda^{(t-1)} + \rho_t\hat{\lambda}.$$

- 8: **until** forever

## Best of Both Worlds

---

**Algorithm 1** Algorithm for hybrid stochastic variational-Gibbs inference.

---

```

for  $t \in 1, \dots, \infty$  do
   $\rho_t \leftarrow \left( \frac{1}{t_0 + t} \right)^\kappa$ 
  sample minibatch  $\mathcal{B}$ 
  for  $d \in \mathcal{B}$  do
    initialize  $\mathbf{z}_d^0$ 
    discard  $B$  burn-in sweeps
    for sample  $s \in 1, \dots, S$  do
      for token  $i \in 1, \dots, N_d$  do
        sample  $z_{di}^s \propto (\alpha + N_{dk}) e^{\mathbb{E}_q[\log \beta_{kw}]}$ 
      end for
    end for
  end for
   $\lambda_{kw}^t \leftarrow (1 - \rho_t) \lambda_{kw}^{t-1} + \rho_t \left( \eta + \frac{D}{|\mathcal{B}|} \hat{N}_{kw} \right)$ 
end for

```

---



## Matching Models and Inference

Zhai and Boyd-Graber, 2013

minibatch-5	minibatch-8	minibatch-10	minibatch-16	minibatch-17	minibatch-39	minibatch-83	minibatch-120
102-club	118-club	132-rock	87-series	82-series	1-annual	0-captain	0-appear
115-issue	128-copy	194-issue	161-issue	162-issue	2-rock	1-appear	1-hulk
127-cover	137-cover	215-series	283-copy	288-copy	3-wolverin	3-hulk	2-wolverin
130-copy	138-issue	217-copy	306-appear	294-appear	4-appear	5-rock	3-annual
197-appear	180-appear	226-cover	307-cover	311-cover	5-comicstrip	6-wolverin	4-copy
289-rock	319-rock	261-appear	502-annual	512-annual	6-series	9-comicstrip	5-rider
450-annual	493-annual	588-annual	814-force	830-force	7-mutant	12-annal	6-comicstrip
584-series	639-series	949-force	1194-rider	4782-wolverin	8-cover	13-mutant	7-cover
811-force	877-force	1074-rider	8944-hulk	9659-hulk	12-issue	15-series	8-force
1090-rider	1003-rider	6038-comicstrip	10819-comicstrip	11527-comicstrip	14-hulk	16-cover	9-captain
	7075-captain	6520-mutant	11301-mutant	12009-mutant	16-copy	19-copy	11-issue
		9569-captain	14335-captain	15040-captain	53-force	23-issue	12-series
					57-rider	280-rider	16-mutant
					86-captain		41-rock
5	8	10	16	17	39	83	
captain	comicstrip	hulk	wolverin	lacy	izzo	gown	
sequitur	mutant	mazelyah	albion				
	patlafountain						

## Matching Models and Inference

Zhai and Boyd-Graber, 2013

minibatch-5	minibatch-8	minibatch-10	minibatch-16	minibatch-17	minibatch-39	minibatch-83	minibatch-120
102-club	118-club	132-rock	87-series	82-series	1-annual	0-captain	0-appear
115-issuee	128-copy	194-issue	161-issue	162-issue	2-rock	1-appear	1-hulk
127-cover	137-cover	215-series	283-copy	288-copy	3-wolverin	3-hulk	2-wolverin
130-copy	138-issue	217-copy	306-appear	294-appear	4-appear	5-rock	3-annual
197-appear	180-appear	226-cover	307-cover	311-cover	5-comicstrip	6-wolverin	4-copy
289-rock	319-rock	261-appear	502-annual	512-annual	6-series	9-comicstrip	5-rider
450-annual	493-annual	588-annual	814-force	830-force	7-mutant	12-annal	6-comicstrip
584-series	639-series	949-force	1194-rider	4782-wolverin	8-cover	13-mutant	7-cover
811-forcee	877-force	1074-rider	8944-hulk	9659-hulk	12-issue	15-series	8-force
1090-rider	1003-rider	6038-comicstrip	10819-comicstrip	11527-comicstrip	14-hulk	16-cover	9-captain
	7075-captain	6520-mutant	11301-mutant	12009-mutant	16-copy	19-copy	11-issue
		9569-captain	14335-captain	15040-captain	53-force	23-issue	12-series
					57-rider	280-rider	16-mutant
					86-captain		41-rock
5	8	10	16	17	39	83	
captain	comicstrip	hulk	wolverin	lacy	izzo	gown	
sequitur	mutant	mazelyah	albion				
	patlafountain						

## Matching Models and Inference

Zhai and Boyd-Graber, 2013

minibatch-5	minibatch-8	minibatch-10	minibatch-16	minibatch-17	minibatch-39	minibatch-83	minibatch-120
102-club	118-club	132-rock	87-series	82-series	1-annual	0-captain	0-appear
115-issuee	128-copy	194-issue	161-issue	162-issue	2-rock	1-appear	1-hulk
127-cover	137-cover	215-series	283-copy	288-copy	3-wolverin	3-hulk	2-wolverin
130-copy	138-issue	217-copy	306-appear	294-appear	4-appear	5-rock	3-annual
197-appear	180-appear	226-cover	307-cover	311-cover	5-comicstrip	6-wolverin	4-copy
289-rock	319-rock	261-appear	502-annual	512-annual	6-series	9-comicstrip	5-rider
450-annual	493-annual	588-annual	814-force	830-force	7-mutant	12-annal	6-comicstrip
584-series	639-series	949-force	1194-rider	4782-wolverin	8-cover	13-mutant	7-cover
811-forcee	877-force	1074-rider	8944-hulk	9659-hulk	12-issue	15-series	8-force
1090-rider	1003-rider	6038-comicstrip	10819-comicstrip	11527-comicstrip	14-hulk	16-cover	9-captain
	7075-captain	6520-mutant	11301-mutant	12009-mutant	16-copy	19-copy	11-issue
		9569-captain	14335-captain	15040-captain	53-force	23-issue	12-series
					57-rider	280-rider	16-mutant
					86-captain		41-rock
5	8	10	16	17	39	83	
captain	comicstrip	hulk	wolverin	lacy	izzo	gown	
sequitur	mutant	mazelyah	albion				
	patlafountain						

## Matching Models and Inference

Zhai and Boyd-Graber, 2013

minibatch-5	minibatch-8	minibatch-10	minibatch-16	minibatch-17	minibatch-39	minibatch-83	minibatch-120
102-club	118-club	132-rock	87-series	82-series	1-annual	0-captain	0-appear
115-issuee	128-copy	194-issue	161-issue	162-issue	2-rock	1-appear	1-hulk
127-cover	137-cover	215-series	283-copy	288-copy	3-wolverin	3-hulk	2-wolverin
130-copy	138-issue	217-copy	306-appear	294-appear	4-appear	5-rock	3-annual
197-appear	180-appear	226-cover	307-cover	311-cover	5-comicstrip	6-wolverin	4-copy
289-rock	319-rock	261-appear	502-annual	512-annual	6-series	9-comicstrip	5-rider
450-annual	493-annual	588-annual	814-force	830-force	7-mutant	12-annal	6-comicstrip
584-series	639-series	949-force	1194-rider	4782-wolverin	8-cover	13-mutant	7-cover
811-forcee	877-force	1074-rider	8944-hulk	9659-hulk	12-issue	15-series	8-force
1090-rider	1003-rider	6038-comicstrip	10819-comicstrip	11527-comicstrip	14-hulk	16-cover	9-captain
	7075-captain	6520-mutant	11301-mutant	12009-mutant	16-copy	19-copy	11-issue
		9569-captain	14335-captain	15040-captain	53-force	23-issue	12-series
					57-rider	280-rider	16-mutant
					86-captain		41-rock
5	8	10	16	17	39	83	
captain	comicstrip	hulk	wolverin	lacy	izzo	gown	
sequitur	mutant	mazelyah	albion				
	patlafountain						

## Matching Models and Inference

Zhai and Boyd-Graber, 2013

minibatch-5	minibatch-8	minibatch-10	minibatch-16	minibatch-17	minibatch-39	minibatch-83	minibatch-120
102-club	118-club	132-rock	87-series	82-series	1-annual	0-captain	0-appear
115-issuee	128-copy	194-issue	161-issue	162-issue	2-rock	1-appear	1-hulk
127-cover	137-cover	215-series	283-copy	288-copy	3-wolverin	3-hulk	2-wolverin
130-copy	138-issue	217-copy	306-appear	294-appear	4-appear	5-rock	3-annual
197-appear	180-appear	226-cover	307-cover	311-cover	5-comicstrip	6-wolverin	4-copy
289-rock	319-rock	261-appear	502-annual	512-annual	6-series	9-comicstrip	5-rider
450-annual	493-annual	588-annual	814-force	830-force	7-mutant	12-annal	6-comicstrip
584-series	639-series	949-force	1194-rider	4782-wolverin	8-cover	13-mutant	7-cover
811-forcee	877-force	1074-rider	8944-hulk	9659-hulk	12-issue	15-series	8-force
1090-rider	1003-rider	6038-comicstrip	10819-comicstrip	11527-comicstrip	14-hulk	16-cover	9-captain
	7075-captain	6520-mutant	11301-mutant	12009-mutant	16-copy	19-copy	11-issue
		9569-captain	14335-captain	15040-captain	53-force	23-issue	12-series
					57-rider	280-rider	16-mutant
					86-captain		41-rock
5	8	10	16	17	39	83	
captain	comicstrip	hulk	wolverin	lacy	izzo	gown	
sequitur	mutant	mazelyah	albion				
	patlafountain						

## Matching Models and Inference

Zhai and Boyd-Graber, 2013

minibatch-5	minibatch-8	minibatch-10	minibatch-16	minibatch-17	minibatch-39	minibatch-83	minibatch-120
102-club	118-club	132-rock	87-series	82-series	1-annual	0-captain	0-appear
115-issuee	128-copy	194-issue	161-issue	162-issue	2-rock	1-appear	1-hulk
127-cover	137-cover	215-series	283-copy	288-copy	3-wolverin	3-hulk	2-wolverin
130-copy	138-issue	217-copy	306-appear	294-appear	4-appear	5-rock	3-annual
197-appear	180-appear	226-cover	307-cover	311-cover	5-comicstrip	6-wolverin	4-copy
289-rock	319-rock	261-appear	502-annual	512-annual	6-series	9-comicstrip	5-rider
450-annual	493-annual	588-annual	814-force	830-force	7-mutant	12-annal	6-comicstrip
584-series	639-series	949-force	1194-rider	4782-wolverin	8-cover	13-mutant	7-cover
811-forcee	877-force	1074-rider	8944-hulk	9659-hulk	12-issue	15-series	8-force
1090-rider	1003-rider	6038-comicstrip	10819-comicstrip	11527-comicstrip	14-hulk	16-cover	9-captain
	7075-captain	6520-mutant	11301-mutant	12009-mutant	16-copy	19-copy	11-issue
		9569-captain	14335-captain	15040-captain	53-force	23-issue	12-series
					57-rider	280-rider	16-mutant
					86-captain		41-rock
5	8	10	16	17	39	83	
captain	comicstrip	hulk	wolverin	lacy	izzo	gown	
sequitur	mutant	mazelyah	albion				
	patlafountain						

## Matching Models and Inference

Zhai and Boyd-Graber, 2013

minibatch-5	minibatch-8	minibatch-10	minibatch-16	minibatch-17	minibatch-39	minibatch-83	minibatch-120
102-club	118-club	132-rock	87-series	82-series	1-annual	0-captain	0-appear
115-issuee	128-copy	194-issue	161-issue	162-issue	2-rock	1-appear	1-hulk
127-cover	137-cover	215-series	283-copy	288-copy	3-wolverin	3-hulk	2-wolverin
130-copy	138-issue	217-copy	306-appear	294-appear	4-appear	5-rock	3-annual
197-appear	180-appear	226-cover	307-cover	311-cover	5-comicstrip	6-wolverin	4-copy
289-rock	319-rock	261-appear	502-annual	512-annual	6-series	9-comicstrip	5-rider
450-annual	493-annual	588-annual	814-force	830-force	7-mutant	12-annal	6-comicstrip
584-series	639-series	949-force	1194-rider	4782-wolverin	8-cover	13-mutant	7-cover
811-forcee	877-force	1074-rider	8944-hulk	9659-hulk	12-issue	15-series	8-force
1090-rider	1003-rider	6038-comicstrip	10819-comicstrip	11527-comicstrip	14-hulk	16-cover	9-captain
	7075-captain	6520-mutant	11301-mutant	12009-mutant	16-copy	19-copy	11-issue
		9569-captain	14335-captain	15040-captain	53-force	23-issue	12-series
					57-rider	280-rider	16-mutant
					86-captain		41-rock
5	8	10	16	17	39	83	
captain	comicstrip	hulk	wolverin	lacy	izzo	gown	
sequitur	mutant	mazelyah	albion				
	patlafountain						

## Matching Models and Inference

Zhai and Boyd-Graber, 2013

minibatch-5	minibatch-8	minibatch-10	minibatch-16	minibatch-17	minibatch-39	minibatch-83	minibatch-120
102-club	118-club	132-rock	87-series	82-series	1-annual	0-captain	0-appear
115-issuue	128-copy	194-issue	161-issue	162-issue	2-rock	1-appear	1-hulk
127-cover	137-cover	215-series	283-copy	288-copy	3-wolverin	3-hulk	2-wolverin
130-copy	138-issue	217-copy	306-appear	294-appear	4-appear	5-rock	3-annual
197-appear	180-appear	226-cover	307-cover	311-cover	5-comicstrip	6-wolverin	4-copy
289-rock	319-rock	261-appear	502-annual	512-annual	6-series	9-comicstrip	5-rider
450-annual	493-annual	588-annual	814-force	830-force	7-mutant	12-annal	6-comicstrip
584-series	639-series	949-force	1194-rider	4782-wolverin	8-cover	13-mutant	7-cover
811-forceee	877-force	1074-rider	8944-hulk	9659-hulk	12-issue	15-series	8-force
1090-rider	1003-rider	6038-comicstrip	10819-comicstrip	11527-comicstrip	14-hulk	16-cover	9-captain
	7075-captain	6520-mutant	11301-mutant	12009-mutant	16-copy	19-copy	11-issue
		9569-captain	14335-captain	15040-captain	53-force	23-issue	12-series
					57-rider	280-rider	16-mutant
					86-captain		41-rock
5	8	10	16	17	39	83	
captain	comicstrip	hulk	wolverin	lacy	izzo	gown	
sequitur	mutant	mazelyah	albion				
	patlafountain						