POLYA-GAMMA TRICK

PG Identity:

$\omega \sim PG(b,0)$ for $b>0$ where $\omega \in \mathbb{R}^+$.

Then $\forall a \in \mathbb{R}$:

$$\boxed{\frac{(e^{\psi})^a}{(1+e^{\psi})^b} = 2^{-b} e^{\kappa\psi} \int_0^{\infty} e^{-\omega\psi^2/2} p(\omega)\, d\omega} \quad \boxed{1}$$

where $\kappa = a - b/2$. Naturally, joint over (ψ,ω):

$$p(\psi,\omega) \propto 2^{-b} e^{\kappa\psi} e^{-\omega\psi^2/2} p(\omega)$$

$\boxed{\text{* Un-Normalized}}$

Conveniently, conditional $\omega|\psi$ is:

$$\boxed{p(\omega|\psi) = \frac{e^{-\omega\psi^2/2} p(\omega)}{\int_0^{\infty} e^{-\omega\psi^2/2} p(\omega)\, d\omega}} \quad \boxed{2}$$

$\sim PG(b,0)$

which is $\sim PG(b,\psi)$!!

Bayesian Logistic Regression

Prior: $\beta \sim \mathcal{N}(\mu, \Sigma)$

Likelihood: $y_i | x_i, \beta \sim \mathrm{Binom}\left(n_i, \frac{1}{1+\exp(\psi_i)}\right)$

$\prod_{\{1,\dots,n_i\}}$ $\psi_i = x_i^T \beta$

n datapts total s.t. $n = \sum_{i=1}^{N} n_i$ where there are N groups of (x_i, y_i)

HOW TO SAMPLE β ?????

AUGMENT WITH PG!!!!!

Likelihood for i^{th} group is:

$$L_i(\beta) = \frac{(e^{x_i^T\beta})^{y_i}}{(1+e^{x_i^T\beta})^{n_i}}$$

$$\propto e^{\kappa_i x_i^T\beta} \int_0^{\infty} e^{-\omega(x_i^T\beta)^2/2} PG(\omega_i | n_i, 0)\, d\omega$$

with $\kappa_i = y_i - n_i/2$. So posterior over β given by:

$$p(\beta|y,\omega) \propto p(\beta) \prod_{i=1}^{N} L_i(\beta|\omega, y) = p(\beta) \prod_{i=1}^{N} e^{\kappa_i x_i^T\beta} e^{-\omega_i (x_i^T\beta)^2/2}$$

$$\propto p(\beta) \prod_{i=1}^{N} \exp\left[\frac{\omega_i}{2}(x_i^T\beta - \kappa_i/\omega_i)^2\right] =$$

$$p(\beta) \exp\left[-\frac{1}{2}(z - X\beta)^T \mathrm{diag}(\omega)(z - X\beta)\right]$$

$z = \left(\dots \frac{\kappa_i}{\omega_i} \dots\right)$!!!!

POLYA-GAMMA CTD...

Pólya-GAMMA GIBBS SAMPLER:

Posterior:

$$\boxed{\beta \mid y, \omega \sim \mathcal{N}(m_\omega, V_\omega)}$$

with mean + covariance

$$V_\omega \; (X^T \mathrm{diag}(\omega) X + \Sigma^{-1})^{-1}$$

with X design matrix

$$m_\omega = V_\omega (X^T K + \Sigma^{-1} \mu)$$

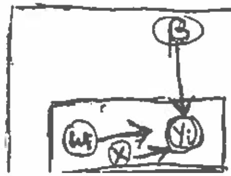
$$K = (y_1 - n_1/2, \ldots, y_N - n_N/2)$$

PG Latent Variables:

$$\boxed{\omega_i \mid \beta \sim PG(n_i, \psi_i)}$$

$$\psi_i = x_i^T \beta$$

ONE ω_i for EACH GROUP



Graphical Model

WHAT IS PG ?

$\omega \sim PG(b,0)$ is an infinite convolution of Gamma distributions

$$\mathbb{E}\{\exp(-\omega t)\} \underset{\text{DEFINITION}}{=} \prod_{k=1}^{\infty} \left(1 + \frac{t}{2\pi^2 (k-1/2)^2}\right)^{-b} \underset{\text{Weierstrass Factorization Theorem}}{=} \frac{1}{\cosh^b(\sqrt{t/2})}$$

Inverting Laplace Transform gives:

$$\omega \overset{D}{=} \frac{1}{2\pi^2} \sum_{k=1}^{\infty} \frac{Ga_k(b,1)}{(k-1/2)^2}$$

$PG(b,c)$ is an <u>exponential tilting</u> of $PG(b,0)$

$$p(\omega \mid b, c) = \frac{\exp\left(\frac{-c^2}{2}\omega\right) p(\omega \mid b, 0)}{\mathbb{E}_{\omega \sim PG(b,0)}\left\{\exp\left(\frac{-c^2}{2}\omega\right)\right\}}$$

Weierstrass Factorization:

$$\omega \overset{D}{=} \frac{1}{2\pi^2} \sum_{k=1}^{\infty} \frac{Ga(b,1)}{(k-1/2)^2 + c^2/4\pi^2}$$

POLYA-GAMMA CTD, CTD...

Fact 1:

Recall

If $\omega \sim PG(b,0)$ then $\boxed{\mathbb{E}(e^{-\omega t}) = \frac{1}{\cosh^b(\psi/2)}}$ and that $\boxed{\cosh z = \frac{1}{2}(e^z + e^{-z})}$

So
$$\frac{(e^\psi)^a}{(1+e^\psi)^b} = \frac{(e^\psi)^a e^{-\frac{\psi}{2}b}}{2^b e^{-\frac{\psi}{2}b} \frac{1}{2^b}(1+e^{2\psi/2})^b}$$

$$= \frac{2^{-b} e^{\kappa\psi}}{\cosh^b(\psi/2)}$$

$$= 2^{-b} e^{\kappa\psi} \mathbb{E}_{\omega}\{\exp(-\omega\psi^2/2)\}$$

$\omega \sim PG(b,0)$ and $\kappa = a - b/2$

✓ Marginalization Identity

Conditional Distribution of $\omega|\psi$ is:

$$p(\omega|\psi) = \frac{e^{-\omega\psi^2/2} \underbrace{p(\omega)}_{PG(b,0)}}{\int_0^\infty e^{-\omega\psi^2/2} p(\omega)\, d\omega}$$

Recall exponential tilting definition of $PG(b,c)$ i.e

$$p(\omega|b,c) = \frac{\exp(-\frac{c^2}{2}\omega)\, p(\omega|b,0)}{\mathbb{E}_{\omega}\{\exp(-c^2/2\,\omega)\}}$$

$\omega \sim PG(b,0)$

Clearly, this $p(\omega|\gamma)$ is also in Polya-Gamma class

✓ Conditional Identity