$\mathcal{N}(M, \Omega, \Sigma)$ - Matrix Variate Gaussian prior

$W \sim \mathcal{N}(M, \Omega, \Sigma)$ $\quad W \in \mathbb{R}^{P \times D}$, D - object dim

M - mean. $\quad$ P - task dim

Ω - row variance (task) $\quad \Omega \in \mathbb{R}^{D \times D}$ $\quad x_n \in \mathbb{R}^D$

Σ_i - column variance (feature) $\quad \Sigma_i \in \mathbb{R}^{P \times P}$ $\quad y_n \in \mathbb{R}^P$

$\Omega = \mathbb{E}\,(W-M)^T(W-M)$ - task

$\Sigma_i = \mathbb{E}\,(W-M)(W-M)$ - column!

$p(w_i) = \int_0^\infty \mathcal{N}(0, \alpha_i^{-1}\Omega_i, \Sigma)$

V - projection from K to P $\quad z_i \in \mathbb{R}^{n \times K}$

$\Sigma_i = \tau I_p$

$\Sigma_{ig} = \sigma^2 I_p$

$Z^T J_L$

$A \sim \mathcal{N}(M_A, \Omega_A, \Sigma_A)$

$B \sim \mathcal{N}(M_B, \Omega_A, \Sigma_B)$

$(A^T B T)_{ij} = a_{ki} b_{km} a_{mj}$ $\quad$ $\mathbb{E}\, a_{ni} b_{kn} a_{nj} = \mathbb{E}\, a_{ki} a_{nj} \cdot \mathbb{E} b_{kn} =$

$A^T B = a_{ik} b_{kj}$ $\quad$ $= (\Omega_{ij}^A \cdot \Sigma_{km}^A + M_{ki} M_{mj}) M_{km}$

$= \Omega_A \cdot tr(\Sigma_A^T M_B) + M_A^T M_B M_B^T$

$(\Omega_{ij} \cdot \Sigma_{km} - M_{ki} M_{mj}) M_{km}$

$(W_i - VZ_i)\Omega_i^{-1}(W_i - VZ_i)^T = W_i \Omega_i^{-1} W_i^T - VZ_i \Omega_i^{-1} W_i - W_i \Omega_i^{-1} (VZ_i)^T +$

$+ VZ_i \Omega_i^{-1} (VZ_i)^T = \Sigma_{wi}\, tr\, \Omega_{wi} \langle \Omega_i^{-1} \rangle + \langle W_i\, Z \Omega_i^{-1} W_i \rangle - \langle V \rangle \langle Z_i \rangle \Omega_i^{-1} \langle W_i \rangle -$

$\mathbb{E}\, a_{in} b_{nm} a_{jn} = \mathbb{E} a_{in} a_{jn} \mathbb{E} b_{nm} = \Sigma_{ij} \cdot \Omega_{nm} M_{nm} + M_{in} M_{jn} \cdot M_{nm} =$

$= \Sigma_A\, tr(\Omega_A M_B) + M_A M_B M_A^T$ $\quad$ $\mathbb{E} a_{in} a_{nj}$

$- \langle W \rangle \langle \Omega_i^{-1} \rangle \langle V \rangle Z_i^T + \mathbb{E}\, V Z_i \Omega_i^{-1} (VZ_i)^T$

$\mathbb{E}\, ABC(AB)^T = \mathbb{E}\, a_{im} b_{mk} c_{ke} \cdot b_{ve} a_{jv} = \mathbb{E} a_{im} a_{jv} \mathbb{E}\, b_{mk} b_{ve} \mathbb{E} c_{ke} =$

$B^T A^T$

$= (\Sigma_{ij}^A \Omega_{mv}^A + M_{im}^A M_{jv}^A)(\Sigma_{mv}^B \Omega_{ke}^B + M_{mk}^B M_{ve}^B) M_{ke}^C =$

$= \Sigma_{ij}^A \Omega_{mv}^A \Sigma_{mv}^B \Omega_{ke}^B M_{ke}^C + M_{im}^A M_{jv}^A \Sigma_{mv}^B \Omega_{ke}^B M_{ke}^C +$

$+ \Sigma_{ij}^A \Omega_{mv}^A \cdot M_{mk}^B M_{ve}^B M_{ke}^C + M_{im}^A M_{jv}^A M_{mk}^B M_{ve}^B M_{ke}^C =$

$= \Sigma_{ij}^A (\Omega_A, \Sigma_B) \cdot (\Omega_B, M_C) + M_A \Sigma_B M_A^T \cdot (\Omega_B, M_C) +$

$+ \Sigma_{ij}^A (\Omega_A, M_B M_C M_B^T) + M_A M_B \cdot M_C (M_A M_B)^T$