1. $f(X) = \det(X^{-1} + A)$, $\nabla_X f$ — ?

$$[D_X f](h) = \langle \nabla_X f, h \rangle = [D_{X^{-1}+A} \det(X^{-1} + A)]([D_X X^{-1}](h))$$

$$[D_Y \det(Y)](h) = \left\langle \left(\frac{\partial \det(Y)}{\partial y_{ij}} \right)_{ij}, h \right\rangle = \left\langle \left(\frac{\partial \left(\sum\limits_k (y_{ik}(-1)^{k+i} M_{ik}) \right)}{\partial y_{ij}} \right)_{ij}, h \right\rangle = \left\langle \left((-1)^{i+j} M_{ij} \right)_{ij}, h \right\rangle =$$

$$= \Big/ Y^{-1} = \frac{1}{\det Y}((-1)^{i+j} M_{ji})_{ij} \Big/ = \left\langle \det Y \cdot Y^{-T}, h \right\rangle$$

$$0 = [D_X I](h) = [D_X(XX^{-1})](h) = X[D_X X^{-1}](h) + [D_X X](h) X^{-1}$$

$$\Longrightarrow [D_X X^{-1}](h) = -X^{-1}[D_X X](h) X^{-1} = -X^{-1} h X^{-1}$$

$$[D_X f](h) = -\left\langle \det Y \cdot Y^{-T}, X^{-1} h X^{-1} \right\rangle = -\det Y \cdot \mathrm{tr}(Y^{-1} X^{-1} h X^{-1}) = -\det Y \cdot \mathrm{tr}(X^{-1} Y^{-1} X^{-1} h) =$$

$$= -\det(X^{-1} + A) \cdot \mathrm{tr}(X^{-1}(X^{-1} + A)^{-1} X^{-1} h)$$

$$\Longrightarrow \nabla_X f = -\det(X^{-1} + A) \cdot \left(X^{-1}(X^{-1} + A)^{-1} X^{-1} \right)^T = -\det(X^{-1} + A) \cdot X^{-T}(X^{-1} + A)^{-T} X^{-T}$$

2. (a) $f(t) = \|(A + tI_n)^{-1} b\|$

$$\frac{\partial}{\partial t}\|(A+tI_n)^{-1}b\| = \frac{b^T(A+tI_n)^{-T}}{\|(A+tI_n)^{-1}b\|}\frac{\partial}{\partial t}(A+tI_n)^{-1}b = -\frac{b^T(A+tI_n)^{-T}}{\|(A+tI_n)^{-1}b\|}(A+tI_n)^{-1}(A+tI_n)^{-1}b$$

$$\frac{\partial^2 f}{\partial t^2} = -\frac{\partial}{\partial t}\frac{b^T(A+tI_n)^{-T}}{\|(A+tI_n)^{-1}b\|}(A+tI_n)^{-1}(A+tI_n)^{-1}b = \frac{\left[b^T(A+tI_n)^{-T}(A+tI_n)^{-1}(A+tI_n)^{-1}b\right]^2}{\|(A+tI_n)^{-1}b\|^3} +$$

$$+\frac{b^T}{\|(A+tI_n)^{-1}b\|}(A+tI_n)^{-T}(A+tI_n)^{-T}(A+tI_n)^{-1}(A+tI_n)^{-1}b +$$

$$+\frac{b^T(A+tI_n)^{-T}}{\|(A+tI_n)^{-1}b\|}(A+tI_n)^{-1}(A+tI_n)^{-1}(A+tI_n)^{-1}b$$

(b) $f(x) = \frac{1}{2}\|xx^T - A\|^2$

$$\frac{\partial}{\partial x}\frac{1}{2}\|xx^T - A\|^2 = \frac{\partial}{\partial x^k}\frac{1}{2}\left((x_i x_j - A_{ij}) \cdot (x^j x^i - A^{ij})\right) =$$

$$= \frac{1}{2}((\delta_i^k x_j + \delta_j^k x_i)(x^j x^i - A^{ji}) + (x_i x_j - A_{ij})(\delta^{kj} x^i + \delta^{ik} x^j)) = 2\|x\|^2 x^T - 2x^T A^T$$

$$\frac{\partial^2 f}{\partial x^l \partial x^k} = 2\delta^{kl} x_j x^j + 4x^k x^l - 2A^{kl} = 2I\|x\|^2 + 4xx^T - 2A$$

3. $\nabla_b L$

$$D_x L[dx] = \langle \nabla_x L; dx \rangle \quad D_b L[db] = \langle \nabla_b L; db \rangle = D_x L[D_b X[db]] = \langle \nabla_x L; A^{-1} db \rangle = \langle A^{-T} \nabla_x L; db \rangle$$

$$\Longrightarrow \nabla_b L = A^{-T} \nabla_x L$$

$\nabla_A L$

$$D_a L[dA] = D_x L[D_A X[dA]] = \langle \nabla_x L; -A^{-1} dA A^{-1} b \rangle = \langle -A^{-T} \nabla_x L; dA A^{-1} b \rangle =$$

$$= \langle -A^{-1} b A^{-T} \nabla_x L; dA \rangle \Longrightarrow \nabla_A L = -A^{-1} b A^{-T} \nabla_x L$$