ECE 641
Fall 2020 MidTerm

Q 2.1

$$p(y|x) = \frac{1}{(2\pi)^{M/2}|R_w|^{M/2}} \exp\left\{-\frac{1}{2}(y-Ax)^T R_w^{-1}(y-Ax)\right\}$$

$$p(x) = \frac{1}{(2\pi)^{N/2}|R_x|^{N/2}} \exp\left\{-\frac{1}{2}x^T R_x^{-1} x\right\}$$

Q 2.2

No, the MLE is not unique.

because $M < N$. This implies that the null space of A has dimension $\geq N - M \geq 1$.

So if $Ax_0 = 0$ for $x_0 \neq 0$, then if x^* is an MLE, then $x^* + \alpha x_0$ is also an MLE.

Q 2.3

No, the MMSE and MAP estimates are the same. This is because both x and y are jointly Gaussian so x given y (i.e., the posterior distribution) is Gaussian. The Gaussian distribution is symmetric and unimodal, so its max = its mean which means MAP = MMSE.

Q2.4

$$\hat{x} = \arg\min_{x} \left\{ \frac{1}{2}(y - Ax) R_w^{-1} (y - Ax) + \frac{1}{2} x^* R_x^{-1} x \right\}$$

$$-A^* R_w^{-1} (y - Ax) + R_x^{-1} x = 0$$

$$-A^* R_w^{-1} y + A^* R_w^{-1} A x + R_x^{-1} x = 0$$

$$(A^* R_w^{-1} A + R_x^{-1}) x = A^* R_w^{-1} y$$

$$\hat{x} = (A^* R_w^{-1} A + R_x^{-1})^{-1} A^* R_w^{-1} y$$

Q 2.5

$$f(x) = \frac{1}{2}(y - Ax)^* R_w^{-1}(y - Ax) + \frac{1}{2} x^* R_x x$$

$$\nabla f(x) = -A^* R_w^{-1}(y - Ax) + R_x x$$

$$x \leftarrow x - \alpha \nabla f$$

$$\leftarrow x + \alpha \left[A^* R_w^{-1}(y - Ax) - R_x x \right]$$

Q2.6

Define $\Lambda = R_w^{-1}$

and $B = R_x^{-1}$

$$f(x) = \log p(x|y) + \text{const}$$

Then

$$= \frac{1}{2}(y - Ax)^t \Lambda (y - Ax) + \frac{1}{2} x^t B x$$

Then $\nabla f(x) = (y - Ax)\Lambda(-A) + x^t B$

$$\nabla \nabla f(x) = A^t \Lambda A + B$$

then

$$\theta_1 \triangleq \frac{\partial f}{\partial x_i} = [\nabla f]_i = -(y - Ax)^T \Lambda A_{*i} + x^t B_{*i}$$

$$\theta_2 \triangleq \frac{\partial^2 f}{\partial x_i^2} = A_{*i}^t \Lambda A_{*i} + B_{ii}$$

$$= \|A_{*i}\|_\Lambda^2 + B_{ii}$$

So then

$$x_i \leftarrow x_i + \frac{-\theta_1}{\theta_2}$$

$$= x_i + \frac{(y - Ax)^T \Lambda A_{*i} - x^T B_{*i}}{\|A_{*i}\|_\Lambda^2 + B_{ii}}$$

Without loss of generality,
$B = B^*$

Q 3.1

$$x^* B + b^* = 0$$

$$\boxed{x = -B^{-1} b}$$

Q 3.2

B^{-1} is N^3 (Gauss-Jordan)

It is computationally expensive because you need to invert a large matrix.

RUBRIC

Q 3.3

Without loss of generality let,

$$\tilde{f}(x; x') = \frac{\alpha}{2}\|x - x'\|^2 + c^T x$$

then pick

$$\alpha = \max\{\lambda_1, \ldots, \lambda_N\}$$

↑ max eigenvalue of B

where

$$B = E \Lambda E^T$$

$$\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_N)$$

and

$$= Bx' + b$$

$$c = \nabla f(x') = B(x') + b$$

So then

$$f(x; x') = \frac{\alpha}{2}\|x\|^2 + \alpha(x')^T x + c^T x$$

$$= \frac{\alpha}{2}\|x\|^2 + (\alpha x' + Bx' + b)^T x$$

Q 3.3

First, define the surrogate function

$$\tilde{f}(x; x') = \frac{\alpha}{2}\|x - x'\|^2 + c'x$$

then

$$\alpha = \lambda_{MAX} = \text{maximum eigenvalue of } B$$

$$c = \nabla f(x') = Bx' + b$$

Then

$$f(x; x') = \tilde{f}(x; x') + \text{const.}$$

$$= \frac{\lambda_{max}}{2}\|x\|^2 \mp \lambda_{max}(x')^t x + c^t x$$

$$= \frac{\lambda_{max}}{2}\|x\|^2 + (Bx' + b \mp \lambda_{max} x')^t x$$

$$f(x; x') = \frac{\lambda_{max}}{2}\|x\|^2 + \left[(B - I\lambda_{max})x' + b\right]^t x$$

Q3.4

$$x'' \leftarrow \arg\min_x f(x; x')$$

$$= \arg\min_x \left\{ \frac{\lambda_{max}}{2} \|x - x'\|^2 + c^T x \right\}$$

Solution

$$\lambda_{max}(x - x') + c \Big|_{x=x''} = 0$$

$$x'' = x' - \frac{c}{\lambda_{max}}$$

$$x' = x' - \frac{Bx' + b}{\lambda_{max}}$$

Algorithm

init $\alpha \leftarrow$ max eigenvalue (B)

init $x \leftarrow 0$

repeat {

$$x \leftarrow x + \frac{1}{\lambda_{max}} (Bx + b)$$

}

Q3.5

Each step requires the evaluation of

$$x \leftarrow x - \frac{1}{2}(Bx + b)$$

NM_0 multiplies per iteration

$O(NM_0)$

$NM_0 \ll N^3$

$\ll$ iterations, N^3 inversion

The Majorization Minimization algorithm requires many fewer multiplies per iteration than direct inversion.

If for example $N = 10^6$ $M_0 = 100$

then

$$10^8 \ll 10^{18}$$

Q 4.1

repeat {

$$(x, v) = \operatorname{argmin} \{ f(x) + h(v) + \frac{\rho}{2}(x - v + u)^2 \}$$

$$u \leftarrow u + (x - v)$$

}

Q4.2

$$F(v) = \underset{x}{\operatorname{argmin}} \left\{ f(x) + \frac{\alpha}{2} \|x - v\|^2 \right\}$$

$$H(v) = \underset{x}{\operatorname{argmin}} \left\{ h(x) + \frac{\alpha}{2} \|x - v\|^2 \right\}$$

$v \leftarrow 0$
$u \leftarrow 0$
repeat {

$x \leftarrow F(v - u)$

$v \leftarrow H(x + u)$

$u \leftarrow u + (x - v)$

}

Q 4.3

It is the same thing

$u \leftarrow 0$
$v \leftarrow 0$
repeat {
$x \leftarrow F(v-u)$
$v \leftarrow \tilde{H}(x+u)$
$u \leftarrow u + (x-v)$
}

Q 4.4

Since

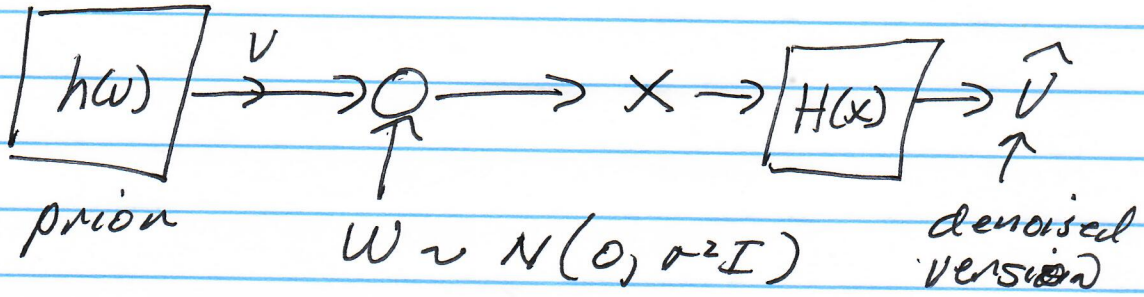
$$H(x) = \arg\min_v \left\{ h(v) + \frac{\alpha}{2}\|v-x\|^2 \right\}$$

define $\sigma^2 = \frac{1}{\alpha}$ then

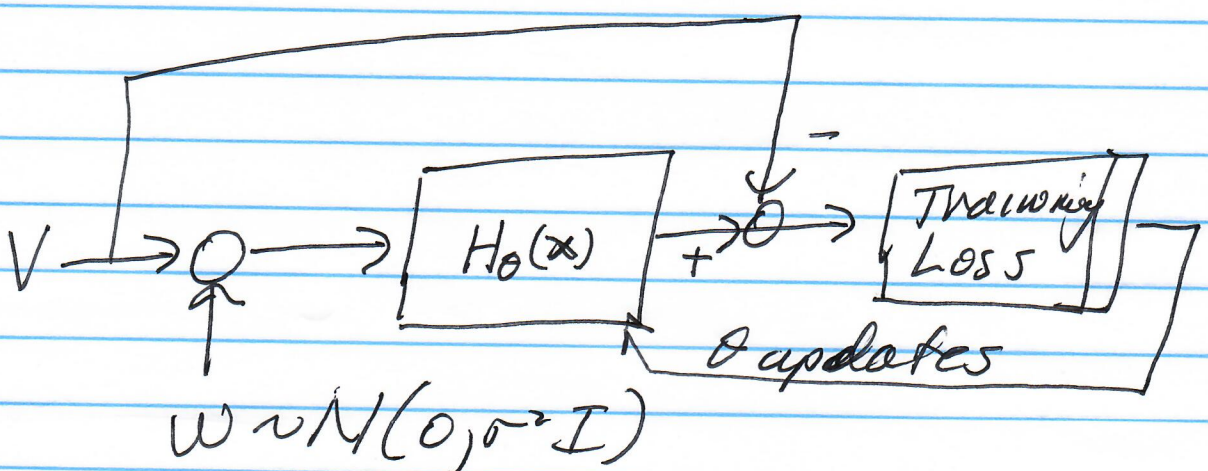
$$= \arg\min_v \left\{ \frac{1}{2\sigma^2}\|x-v\|^2 + h(v) \right\}$$

↑ MAP denoiser for

So the virtual model is



Training



Q5

This mean that you must have a goal first. Then you can design tactics to achieve the goal (i.e. strategy).

So in research, first determine the problem you would like to solve. Then come up with a method that will solve the problem.