Задача 1.

$E_{x,y}[(y-\mu(x))^2] = E_{x,y}[(y-\mu(x)+E(y|x)-E(y|x))] = E_{x,y}[((y-E(y|x))+(E(y|x)-\mu(x)))^2] =$

$= E_{x,y}[(y-E(y|x))^2] + E_{x,y}[(E(y|x)-\mu(x))^2] + 2E_{x,y}[(y-E(y|x))(E(y|x)-\mu(x))]$

$E_{x,y}[(y-E(y|x))(E(y|x)-\mu(x))] = \int_X \int_Y (y-E(y|x))(E(y|x)-\mu(x))p(x,y)dxdy =$

$= \int_X (E(y|x)-\mu(x)) \left[\int_Y (y-E(y|x))p(x,y)dy\right] dx$

$\int_Y (y-E(y|x))p(x,y)dy = \int_Y yp(y|x)p(x)dy - \int_Y E(y|x)p(x,y)dy = p(x)\int_Y yp(y|x)dy - E(y|x)\int_Y p(x,y)dy =$

$= p(x)E(y|x) - E(y|x)p(x) = 0$, $\Rightarrow \mathrm{E}_{x,y}[(y-\mu(x))^2] = E_{x,y}[(y-E(y|x))^2] + E_{x,y}[(E(y|x)-\mu(x))^2]$

$L(\mu) = E_{X^\ell,y^\ell}[E_{x,y}(y-E(y|x))^2] + E_{X^\ell,y^\ell}[E_{x,y}(E(y|x)-\mu(x))^2]$, где первое слагаемое - шумовая компонента, которая не зависит от X^ℓ, перепишем:

$L(\mu) = E_{x,y}(y-E(y|x))^2 + E_{X^\ell,y^\ell}[E_{x,y}(E(y|x)-\mu(x))^2]$

$E_{X^\ell,y^\ell}[E_{x,y}(E(y|x)-\mu(x))^2] = E_{X^\ell,y^\ell}[E_{x,y}(E(y|x)-\mu(x)+E_{X^\ell,y^\ell}(\mu(x))-E_{X^\ell,y^\ell}(\mu(x)))^2] =$

$= E_{x,y}[E_{X^\ell,y^\ell}((E(y|x)-E_{X^\ell,y^\ell}(\mu(x))) + (E_{X^\ell,y^\ell}(\mu(x))-\mu(x)))^2] =$

$= E_{x,y}[E_{X^\ell,y^\ell}(E(y|x)-E_{X^\ell,y^\ell}(\mu(x)))^2] + E_{x,y}[E_{X^\ell,y^\ell}(E_{X^\ell,y^\ell}(\mu(x))-\mu(x))^2] +$

$+2E_{x,y}[E_{X^\ell,y^\ell}(E(y|x)-E_{X^\ell,y^\ell}(\mu(x)))(E_{X^\ell,y^\ell}(\mu(x))-\mu(x))]$,

первое слагаемое можем перписать как $E_{x,y}(E(y|x)-E_{X^\ell,y^\ell}(\mu(x)))^2$

Рассмотрим последнее слагаемое: $E_{x,y}[E_{X^\ell,y^\ell}(E(y|x)-E_{X^\ell,y^\ell}(\mu(x)))(E_{X^\ell,y^\ell}(\mu(x))-\mu(x))] =$

$= E_{x,y}[(E(y|x)-E_{X^\ell,y^\ell}(\mu(x)))E_{X^\ell,y^\ell}(E_{X^\ell,y^\ell}(\mu(x))-\mu(x))] =$

$= [E_{X^\ell,y^\ell}(E_{X^\ell,y^\ell}(\mu(x))-\mu(x)) = E_{X^\ell,y^\ell}(\mu(x)) - E_{X^\ell,y^\ell}(\mu(x)) = 0] = 0$

$\Rightarrow \mathrm{E}_{X^\ell,y^\ell}[E_{x,y}(E(y|x)-\mu(x))^2] = E_{x,y}(E(y|x)-E_{X^\ell,y^\ell}(\mu(x)))^2 + E_{x,y}[E_{X^\ell,y^\ell}(E_{X^\ell,y^\ell}(\mu(x))-\mu(x))^2]$

$\Rightarrow \mathrm{L}(\mu) = E_{X^\ell,y^\ell}[E_{x,y}[(y-\mu(x))^2]] = E_{x,y}(y-E(y|x))^2 + E_{x,y}(E(y|x)-E_{X^\ell,y^\ell}(\mu(x)))^2 +$

$+E_{x,y}[E_{X^\ell,y^\ell}(E_{X^\ell,y^\ell}(\mu(x))-\mu(x))^2]$

Задача 5.

Нам необходимо уменьшить сложность вычисления последнего слагаемого.

Заметим, что $\langle\sum_{i=1}^d x_i v_i, \sum_{i=1}^d x_i v_i\rangle = \sum_{i=1}^d \sum_{j=i+1}^d x_i x_j \langle v_i, v_j\rangle + \sum_{i=1}^d x_i^2 \langle v_i, v_i\rangle + \sum_{i=1}^d \sum_{j=1}^{i-1} x_i x_j \langle v_i, v_j\rangle$,

Заметим также, что $\sum_{i=1}^d \sum_{j=i+1}^d x_i x_j \langle v_i, v_j\rangle = \sum_{i=1}^d \sum_{j=1}^{i-1} x_i x_j \langle v_i, v_j\rangle$, тогда:

$\langle\sum_{i=1}^d x_i v_i, \sum_{i=1}^d x_i v_i\rangle = 2\sum_{i=1}^d \sum_{j=i+1}^d x_i x_j \langle v_i, v_j\rangle + \sum_{i=1}^d x_i^2 \langle v_i, v_i\rangle$

Тогда $\sum_{i=1}^d \sum_{j=i+1}^d x_i x_j \langle v_i, v_j\rangle = \frac{1}{2}(\langle\sum_{i=1}^d x_i v_i, \sum_{i=1}^d x_i v_i\rangle - \sum_{i=1}^d x_i^2 \langle v_i, v_i\rangle)$

Оба члена получившейся разности считаются линейно, подставив в исходную сумму, также получим линейную сложность.

Задача 6.

У метода Дениса основное преимущество - дешевизна. Этим методом можно пользоваться, не имея дорогих серверов и большого количества уже пользующихся сервисом пользователей. По сути, для работы этого метода достаточно иметь собственно алгоритм, который будет предсказывать и достаточное количество данных, чтобы сделать какие-то выводы. С другой стороны, нельзя точно гарантировать что алгоритм работает хорошо, если он хорошо себя показал на тестовых данных. Возможно, тестовые данные слишком старые и не показывают реальных потребностей пользователей в данный момент, возможно, то, что было модным пару месяцев назад, когда были собраны данные, больше никто не покупает.

Наоборот, метод Андрея очень дорогой, для него понадобится продолжительное время работающий магазин с большой клиентской базой, но при этом с его помощью можно оценить, как действительно работает алгоритм "в бою".

Мне кажется, что нужно использовать оба метода - сначала использовать метод Дениса, чтобы получить какую-то работающую рекомендательную систему, затем, когда наберется достаточно клиентов, начать проводить AB-тестирование и fine-tune'ить рекомендательную систему.