Austria I i). We can use a categorial multinoully distribution to describe this scenario. We will have six parameters: M. M. M. M. M. M. the probabilities for side 1, 2, 3, 4, 5, 16 comes up 2). If we have a fair dice then M. = M. = M. = M. = M. = 5 3) if the die always rolls two then
Mz=1, M= M3/= My= M3 = Mb= D. 4) the domain of powometers is [D,1] Inc [a,1] Mzc [o,1] Mc E [o,1] MUE [o,1] VSE [o,1] Mb E [o,1]

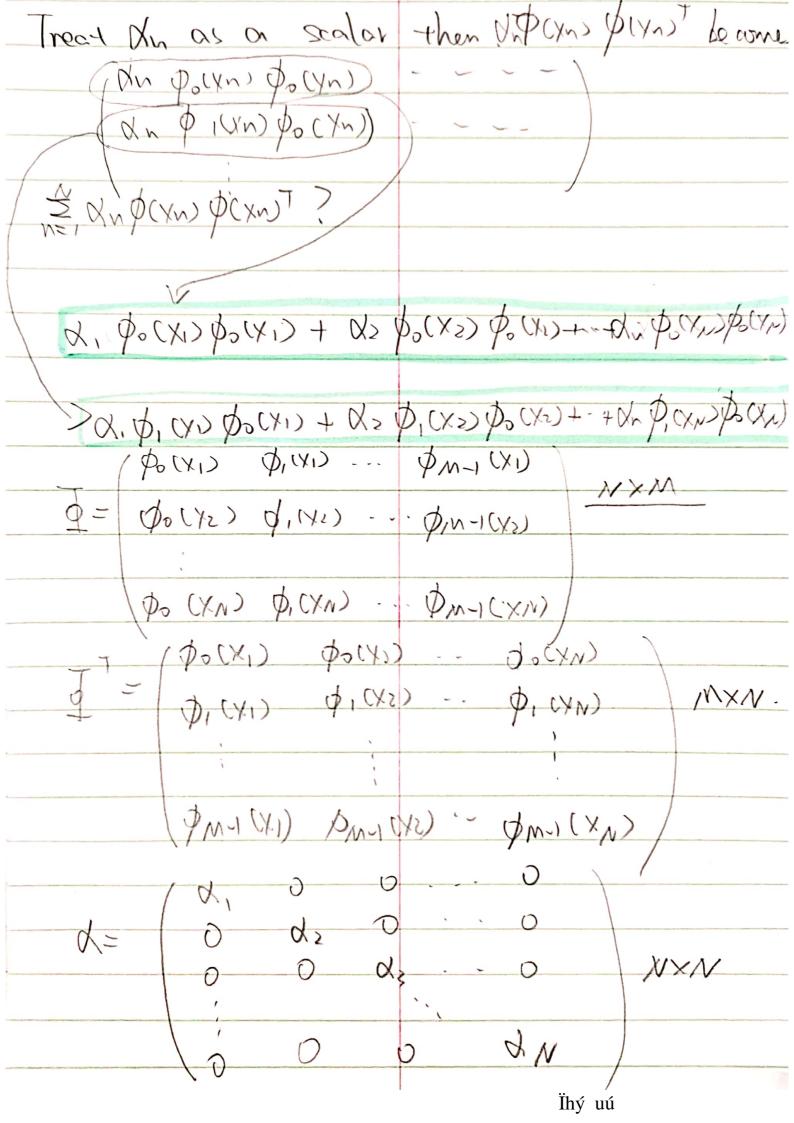
Question 2 Es(w)= \frac{2}{2} \text{Xn stn-wTpCxn}g = \$: (th - W) p(xn) (-p(xn)) . dn TEDOWN = E : (tndn - wtdn p(xn)) (- p(xn)) T D=[0,0,0]

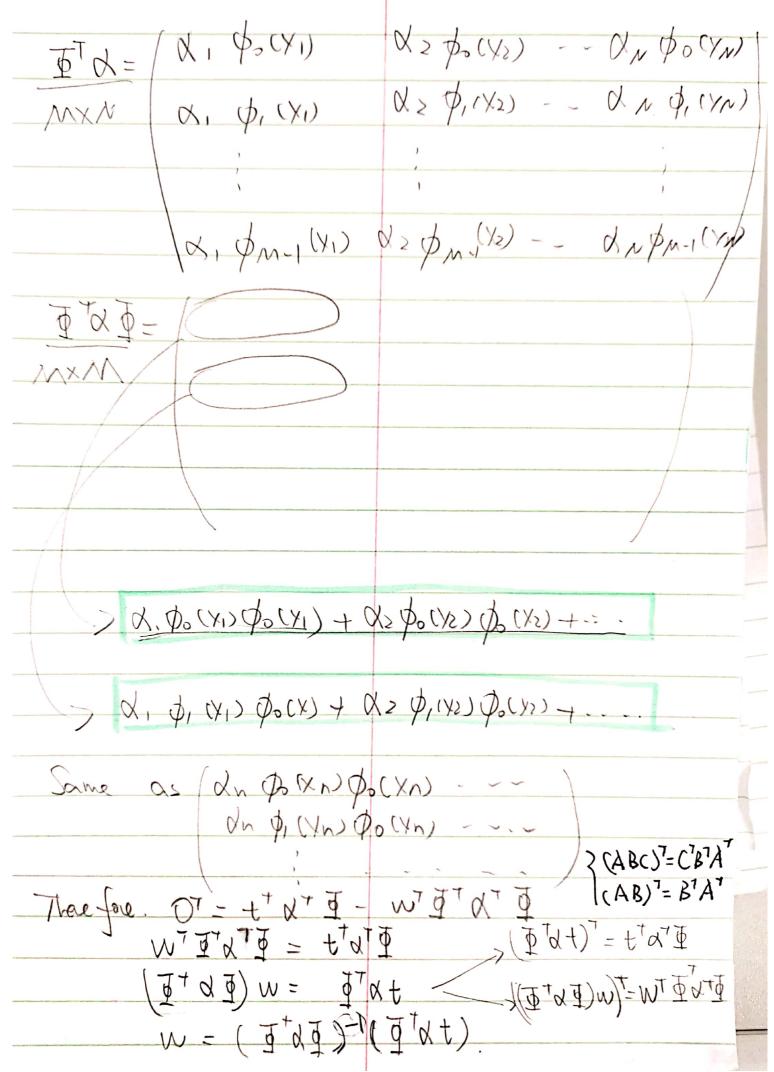
D=[0,0,0]

D=[0,0,0]

D=[0,0,0] [(ux fm Q] Set the gradient to 0 0= DES(w) = = (tndn-w/dn p(xn))(-p(xn)) $D' = \sum_{n=1}^{\infty} -t_n \alpha_n \beta(x_n) + W' \sum_{n=1}^{\infty} d_n \beta(y_n) \beta(x_n)$ 0= ++ 0+ = - m+ 1 x 1 Why? Whis part as an example: Pr (yn) [JMFYN] Jockes Dockes Dockes Dockes かいからいか Q, (yn) (700yn)

Ïhý uú

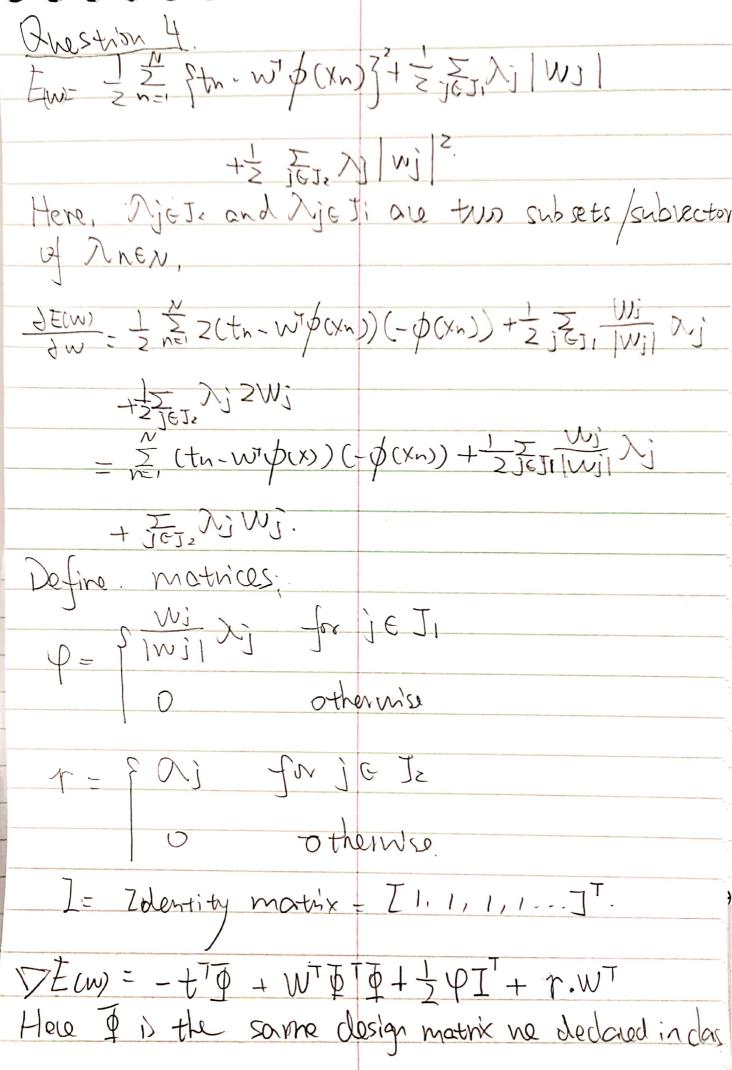




Ïhý uú

Questim 3 E(\w)= = = = = = ||w||^2. RMS: EPMS = NZE(W* >/N D. No. The training set and vahidation set are both randomly distributed data, from the dataser. The is no growntees on the relationship between training orner and validation error 27 the model is overfit the validation error is probably higher. than the training error Good fit: Validarin error hur, slight higher than the training orm. hukum fit. Valid atin error low, training error high Under Sig: Validation error and training error both trafi Generally speaking, training error will almost always underestimate the validation error, But it is possible for the would ation error to be loss than the training, 2) Yes. Degree to polynamial contains Degree 9 polynomial. The unregularized regression gies us the optimal solution which means Degree to polymonial mostly fits the data better. In the worst case, the training error

for them two are equal. Buy if we change training error to testing error for this question then the answer show be "No". In most cases the testing error for regularized regression & lower that hurregularized regression since the degree=20 is very high and is highly likely to Cause overfitting.
But this D not garmanteed. If we got a reak model than the regularized power even more and make the testing error largor compared with the un regularized one



Ïhý uú

Questin 5.
5.1) Niger. 313.7/1000 = 31.37%
Sierra Leone. 185.3/1000 - 18.53%
na. values = "" will set the missing
fectures to NA values.
Then we will use nonmean() to find the overage value for each column / feature,
M, where (np. isnan) will find the
copydinance in dies for the NA value.
and ne ossign the allrage value to
the missing parts according to their
NB: np. where (np. i= nan) nill return
the now and colum indies Bry
np. take (mean-vails, inds[1]) > ue
only need the column indies (the the
only need the column indies since the onerange is an XIX 40 moths/vector.
) X X /