**MATH-564**

**HOMEWORK – 4**

**PART-I**

1. Obtain the variance inflation factors. Are there any indications that serious multi-collinearity problems are present?

**SOLUTION:-**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| x1 | x2 | x3 | x4 | x5 | x6 |
| 1.173570 | 13.276326 | 12.614763 | 2.048616 | 1.397569 | 2.050460 |

2 variables, **variable 8 and variable 9** have serious multi-collinearity issues at present.

1. Obtain the studentized deleted residuals and prepare a dot plot of these residuals. Are there any outliers present?

**SOLUTION:-**

****

**There are 9 outliers.**

1. Obtain the diagonal elements of the hat matrix. Using the rule of thumb (if hii > 3h, then consider xi extreme), identify any extreme X values.

**SOLUTION:-**

DIAGONAL ELEMENTS OF THE HAT MATRIX:-

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 0.45101833 | 0.19479252 | 0.07792132 | 0.03229018 | 0.02857239 | 0.0517693 | 0.01557477 | 0.1119271 | 0.0289475 |
| 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 0.01869702 | 0.04660584 | 0.05017841 | 0.02624007 | 0.00669558 | 0.02511484 | 0.04320644 | 0.03365274 | 0.01144516 |
| 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 |
| 0.05868081 | 0.00948 | 0.03834005 | 0.01246452 | 0.01393633 | 0.01665408 | 0.02915636 | 0.00507552 | 0.01502996 |
| 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 |
| 0.0084632 | 0.00999047 | 0.01020297 | 0.00963686 | 0.05542151 | 0.01089177 | 0.02411954 | 0.01075517 | 0.0368692 |
| 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 |
| 0.0111836 | 0.00485881 | 0.03501601 | 0.0216281 | 0.03999115 | 0.03270616 | 0.01900674 | 0.01050013 | 0.00993038 |
| 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 |
| 0.01329615 | 0.02446197 | 0.1058448 | 0.00481115 | 0.03144532 | 0.00945106 | 0.01664592 | 0.05184661 | 0.01384644 |
| 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 |
| 0.00620529 | 0.01450691 | 0.00973096 | 0.0261323 | 0.01310638 | 0.00437914 | 0.01009522 | 0.01144729 | 0.00495981 |
| 64 | 65 | 66 | 67 | 68 | 69 | 70 | 71 | 72 |
| 0.02141179 | 0.01561915 | 0.00633642 | 0.04592647 | 0.01521748 | 0.01995553 | 0.03307145 | 0.0131781 | 0.02364857 |
| 73 | 74 | 75 | 76 | 77 | 78 | 79 | 80 | 81 |
| 0.0216348 | 0.0072158 | 0.00734764 | 0.02366195 | 0.00407575 | 0.00649869 | 0.01463073 | 0.01039 | 0.01639682 |
| 82 | 83 | 84 | 85 | 86 | 87 | 88 | 89 | 90 |
| 0.01301823 | 0.00613792 | 0.0069575 | 0.01456398 | 0.00823321 | 0.01607206 | 0.0043236 | 0.00738445 | 0.01030046 |
| 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 |
| 0.02340361 | 0.00687023 | 0.00327536 | 0.00420856 | 0.06296083 | 0.01850889 | 0.00903377 | 0.01568353 | 0.00682549 |
| 100 | 101 | 102 | 103 | 104 | 105 | 106 | 107 | 108 |
| 0.01066286 | 0.0037986 | 0.01372716 | 0.0048154 | 0.00948459 | 0.00764656 | 0.01112065 | 0.005253 | 0.00780356 |
| 109 | 110 | 111 | 112 | 113 | 114 | 115 | 116 | 117 |
| 0.00613514 | 0.00762415 | 0.01555274 | 0.0121657 | 0.02007463 | 0.00656768 | 0.01021536 | 0.00539293 | 0.03216449 |
| 118 | 119 | 120 | 121 | 122 | 123 | 124 | 125 | 126 |
| 0.01214339 | 0.01148353 | 0.00587941 | 0.00496914 | 0.00462834 | 0.04740126 | 0.00735398 | 0.01324286 | 0.00697767 |
| 127 | 128 | 129 | 130 | 131 | 132 | 133 | 134 | 135 |
| 0.00574931 | 0.11240853 | 0.01348641 | 0.01273365 | 0.00997921 | 0.00803978 | 0.03119986 | 0.00697857 | 0.00739697 |
| 136 | 137 | 138 | 139 | 140 | 141 | 142 | 143 | 144 |
| 0.00557329 | 0.0133645 | 0.00405887 | 0.00681368 | 0.02006801 | 0.0133462 | 0.00944913 | 0.00546122 | 0.00414345 |
| 145 | 146 | 147 | 148 | 149 | 150 | 151 | 152 | 153 |
| 0.00582643 | 0.00566284 | 0.00479167 | 0.01025602 | 0.00760045 | 0.0103933 | 0.00820687 | 0.01491121 | 0.00466245 |
| 154 | 155 | 156 | 157 | 158 | 159 | 160 | 161 | 162 |
| 0.00881928 | 0.05234344 | 0.00581553 | 0.00363496 | 0.0107291 | 0.00483116 | 0.01097268 | 0.00507299 | 0.00435095 |
| 163 | 164 | 165 | 166 | 167 | 168 | 169 | 170 | 171 |
| 0.01143629 | 0.00558196 | 0.00523023 | 0.00873874 | 0.00487009 | 0.02519034 | 0.0048617 | 0.01524963 | 0.02266096 |
| 172 | 173 | 174 | 175 | 176 | 177 | 178 | 179 | 180 |
| 0.01231774 | 0.02547736 | 0.00438633 | 0.00663332 | 0.01664863 | 0.01662156 | 0.00751961 | 0.00489664 | 0.01450621 |
| 181 | 182 | 183 | 184 | 185 | 186 | 187 | 188 | 189 |
| 0.00679856 | 0.00922929 | 0.00703053 | 0.01765177 | 0.0063165 | 0.00755933 | 0.02985137 | 0.08309756 | 0.00577355 |
| 190 | 191 | 192 | 193 | 194 | 195 | 196 | 197 | 198 |
| 0.00649617 | 0.00468351 | 0.01224719 | 0.00383348 | 0.0033041 | 0.00504819 | 0.02120225 | 0.00575965 | 0.00405181 |
| 199 | 200 | 201 | 202 | 203 | 204 | 205 | 206 | 207 |
| 0.01046994 | 0.00405676 | 0.03973095 | 0.01425483 | 0.00484192 | 0.00626604 | 0.00400259 | 0.07497916 | 0.00583984 |
| 208 | 209 | 210 | 211 | 212 | 213 | 214 | 215 | 216 |
| 0.00839035 | 0.00701537 | 0.00635839 | 0.00562688 | 0.00816215 | 0.0109426 | 0.01012208 | 0.01513601 | 0.00453043 |
| 217 | 218 | 219 | 220 | 221 | 222 | 223 | 224 | 225 |
| 0.01003095 | 0.0052431 | 0.00630219 | 0.00482224 | 0.00504717 | 0.01284954 | 0.00931294 | 0.00613138 | 0.01352792 |
| 226 | 227 | 228 | 229 | 230 | 231 | 232 | 233 | 234 |
| 0.00782925 | 0.00347004 | 0.00909297 | 0.01228736 | 0.00913863 | 0.00954778 | 0.0070578 | 0.00610671 | 0.006307 |
| 235 | 236 | 237 | 238 | 239 | 240 | 241 | 242 | 243 |
| 0.00792285 | 0.00795365 | 0.01385456 | 0.00458181 | 0.01735527 | 0.00337467 | 0.01143344 | 0.00534209 | 0.0110557 |
| 244 | 245 | 246 | 247 | 248 | 249 | 250 | 251 | 252 |
| 0.0073346 | 0.01546835 | 0.0187916 | 0.00701335 | 0.01657465 | 0.02207025 | 0.00732175 | 0.00904822 | 0.00888177 |
| 253 | 254 | 255 | 256 | 257 | 258 | 259 | 260 | 261 |
| 0.00461093 | 0.00624653 | 0.00572051 | 0.00665607 | 0.00657443 | 0.01435225 | 0.0268088 | 0.00519631 | 0.01309119 |
| 262 | 263 | 264 | 265 | 266 | 267 | 268 | 269 | 270 |
| 0.0324376 | 0.00633865 | 0.00427307 | 0.00353975 | 0.00844035 | 0.0122725 | 0.01285946 | 0.02660418 | 0.00716421 |
| 271 | 272 | 273 | 274 | 275 | 276 | 277 | 278 | 279 |
| 0.01028177 | 0.04546927 | 0.00788865 | 0.00860574 | 0.00654958 | 0.00910953 | 0.00387384 | 0.00537736 | 0.01636344 |
| 280 | 281 | 282 | 283 | 284 | 285 | 286 | 287 | 288 |
| 0.00648213 | 0.00946867 | 0.00648719 | 0.00598061 | 0.00561846 | 0.01003634 | 0.00573462 | 0.01503641 | 0.00757625 |
| 289 | 290 | 291 | 292 | 293 | 294 | 295 | 296 | 297 |
| 0.00673724 | 0.00408155 | 0.00702907 | 0.01182107 | 0.01513138 | 0.01794426 | 0.00408102 | 0.00450921 | 0.00643271 |
| 298 | 299 | 300 | 301 | 302 | 303 | 304 | 305 | 306 |
| 0.00961693 | 0.00582606 | 0.00499969 | 0.03320767 | 0.02123723 | 0.07452592 | 0.00843414 | 0.00461315 | 0.00471476 |
| 307 | 308 | 309 | 310 | 311 | 312 | 313 | 314 | 315 |
| 0.00532086 | 0.01112678 | 0.0051264 | 0.02286907 | 0.01010316 | 0.00672762 | 0.01595567 | 0.01771007 | 0.00873498 |
| 316 | 317 | 318 | 319 | 320 | 321 | 322 | 323 | 324 |
| 0.00789446 | 0.00471755 | 0.00542022 | 0.01198793 | 0.0087341 | 0.00794083 | 0.02042649 | 0.00408856 | 0.00521739 |
| 325 | 326 | 327 | 328 | 329 | 330 | 331 | 332 | 333 |
| 0.01538851 | 0.00746404 | 0.00604808 | 0.01097909 | 0.00878226 | 0.01158264 | 0.01016041 | 0.00399831 | 0.00542909 |
| 334 | 335 | 336 | 337 | 338 | 339 | 340 | 341 | 342 |
| 0.0244251 | 0.00839856 | 0.01577847 | 0.08219586 | 0.00774127 | 0.01113086 | 0.01746632 | 0.00441593 | 0.01117701 |
| 343 | 344 | 345 | 346 | 347 | 348 | 349 | 350 | 351 |
| 0.00454383 | 0.03177275 | 0.00820728 | 0.01475895 | 0.00851357 | 0.01266983 | 0.00510684 | 0.00345514 | 0.00534978 |
| 352 | 353 | 354 | 355 | 356 | 357 | 358 | 359 | 360 |
| 0.00484717 | 0.00761117 | 0.00809325 | 0.00810364 | 0.01106561 | 0.04642174 | 0.00599872 | 0.00714857 | 0.01674713 |
| 361 | 362 | 363 | 364 | 365 | 366 | 367 | 368 | 369 |
| 0.00530508 | 0.00773795 | 0.06549221 | 0.01331986 | 0.00794218 | 0.00643294 | 0.00858945 | 0.01878151 | 0.00347974 |
| 370 | 371 | 372 | 373 | 374 | 375 | 376 | 377 | 378 |
| 0.00665487 | 0.00610022 | 0.00487268 | 0.00639734 | 0.01390459 | 0.01048427 | 0.00589703 | 0.00818001 | 0.00644519 |
| 379 | 380 | 381 | 382 | 383 | 384 | 385 | 386 | 387 |
| 0.00646964 | 0.02123557 | 0.00570468 | 0.00436379 | 0.01010161 | 0.00834782 | 0.0101467 | 0.01379072 | 0.00911121 |
| 388 | 389 | 390 | 391 | 392 | 393 | 394 | 395 | 396 |
| 0.00609164 | 0.01061228 | 0.01200306 | 0.00483335 | 0.02403443 | 0.01194827 | 0.00611449 | 0.01130858 | 0.05687733 |
| 397 | 398 | 399 | 400 | 401 | 402 | 403 | 404 | 405 |
| 0.00632215 | 0.02679815 | 0.01953488 | 0.00926115 | 0.00837899 | 0.00798453 | 0.0038538 | 0.10267529 | 0.04447009 |
| 406 | 407 | 408 | 409 | 410 | 411 | 412 | 413 | 414 |
| 0.01710986 | 0.0139316 | 0.00682293 | 0.01665139 | 0.02913015 | 0.00502592 | 0.03058438 | 0.00524422 | 0.00912605 |
| 415 | 416 | 417 | 418 | 419 | 420 | 421 | 422 | 423 |
| 0.06238255 | 0.00885544 | 0.01115599 | 0.01184423 | 0.00659146 | 0.0086478 | 0.00838656 | 0.02303077 | 0.00673128 |
| 424 | 425 | 426 | 427 | 428 | 429 | 430 | 431 | 432 |
| 0.00707179 | 0.01261344 | 0.00732117 | 0.00699061 | 0.00590938 | 0.0141239 | 0.01482569 | 0.01608651 | 0.00949671 |
| 433 | 434 | 435 | 436 | 437 | 438 | 439 | 440 |  |
| 0.02938715 | 0.00545229 | 0.0062404 | 0.03007406 | 0.0301264 | 0.01546922 | 0.01096899 | 0.00383084 |  |
|  |  |  |  |  |  |  |  |  |

EXTREME X VALUES:-

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 2 | 3 | 6 | 8 | 12 | 19 | 32 | 48 | 53 |
| 0.45101832 | 0.19479252 | 0.07792132 | 0.0517693 | 0.1119271 | 0.05017841 | 0.05868081 | 0.05542151 | 0.1058448 | 0.05184661 |
| 95 | 128 | 155 | 188 | 206 | 303 | 337 | 363 | 396 | 404 |
| 0.06296083 | 0.11240853 | 0.05234344 | 0.08309756 | 0.07497916 | 0.07452592 | 0.08219586 | 0.06549221 | 0.05687733 | 0.10267529 |
| 415 |  |  |  |  |  |  |  |  |  |
| 0.06238255 |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |

1. Cases 2, 8, 48, 128, 206, and 404 are outlying with respect to their X values, and cases 2 and 6 are outlying with respect to their Y values. Obtain DFFITS, DFBETAS and Cook’s distance values for these cases to assess their influence. What do you conclude?

**SOLUTION:-**

We check the threshold for DFFITS (by ols\_plot\_dffits(mod)), which is 0.25 and based on that we get influence values which is **6**. So only **6** has got values which is greater than 0.25.

2 6 8 48 128 206

-0.57559491 6.92176520 0.02167814 0.04879748 0.07656083 0.24389295

404

-0.06173662

We check the threshold for DFBETAS (by ols\_plot\_dfbetas(mod)), which is 0.1 and based on that we get influence values which is **2, 6**. So, only **2 and 6** has got values which is greater than 0.1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Intercept | x1 | x2 | x3 | x4 | x5 | x6 |  |
| 2 | -0.0565987 | 0.01079253 | 0.07103094 | -0.2412882 | 0.14224772 | -0.0291515 | 0.10359102 |  |
| 6 | -1.712836 | 0.51275988 | -4.0025622 | 5.0134918 | 1.43936412 | 0.71459504 | 1.42605415 |  |
| 8 | -0.005598 | 0.00271008 | -0.0178494 | 0.02000975 | -0.0001379 | 0.00327588 | 0.00495761 |  |
| 48 | 0.00428637 | -0.0091892 | 0.04290249 | -0.0426039 | 0.01010578 | -0.0069171 | 0.0042799 |  |
| 128 | -0.0268677 | 0.00036465 | 0.01208476 | -0.0179494 | 0.05367045 | 0.02510301 | 0.01979965 |  |
| 206 | -0.1101023 | -0.0282703 | -0.0229477 | -0.0073253 | 0.11338631 | -0.0239341 | 0.22196234 |  |
| 404 | 0.02697068 | -0.0101053 | -0.0055056 | 0.00858268 | -0.0071767 | -0.0534785 | -0.0121107 |  |

We know the threshold which is p = 0.5 and based in that we get influence values which is **6**. So only variable **6** has pf>0.5.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Cook:- |  |  |  |  |  |  |
| 2 | 6 | 8 | 48 | 128 | 206 | 404 |
| 4.73E-02 | 2.26E+00 | 6.73E-05 | 3.41E-04 | 8.39E-04 | 8.50E-03 | 5.46E-04 |
| pf(cook,7,433) |  |  |  |  |  |  |
| 2 | 6 | 8 | 48 | 128 | 206 | 404 |
| 1.42E-04 | 9.71E-01 | 1.76E-14 | 5.14E-12 | 1.20E-10 | 3.90E-07 | 2.67E-11 |

**PART-II**

First, we start with only intercept. Then we run the add1() function with which we see the p values for all the predictors. We need to take the predictors which have p values less than alpha-to-enter i.e., less than 0.15.

Now, all the predictors have p-values less than 0.15. So we select the predictor with lowest p-value, which is x4 whose p-value is 0.0005762. Therefore x4 enters the model.

Now we run add1() function again and see which predictor to be selected. We see x2 has got p-value greater than 0.15. Therefore we reject it. Now, x1 and x3 has got p-values less than 0.15, so we see which has got the lowest value. We see x1 has got the lowest p-value, i.e., 1.105e-06. Therefore x1 enters the model. We see the summary of the model and get to know that x4 and x1 has got p-values less than 0.15. Therefore model is valid.

Now we run add1() function again and see which predictor to be selected. We see both x2 and x3 has p-values less than 0.15 and x2 has got the lowest. Therefore we add x2 into the model. Now, we check the summary of the model and realize that p-value of x4 has increased. Therefore we drop x4 from the model.

Now we run add1() function again and see which predictor to be selected. We see that both x4 and x3 has got p-values greater than 0.15. Therefore, we stop here.

Thus our final model will have **x1 and x2**.