Index

References to figures are given in italic type. References to tables are given in bold type.

adaptive partial least squares, 287–91
see also moving-window principal component analysis; time-varying processes
alternative hypotheses, 13
application delay, 253–4, 261–5, 275–8, 276–8, 282–6
bootstrapping, 77–9, 287
broken stick model, 70
canonical correlation analysis (CCA), 408–9
catalyst fluidization, 139–40
central limit theorem, 128, 208, 294, 341–4, 367–8
central moments, 312, 315
chemical reaction process example, 123–4
correlation matrix construction, 126
covariance matrix, 126
fault diagnosis, 134–40
temperature variables, 132–3, 140
variable reconstruction, 126–7
Cholesky decomposition, 168, 177–82
column pressure, 225–6
column temperature, 157
common cause variation, 11–12, 30–2, 40–1, 49–52, 55, 58–9, 64
see also t-score
computational efficiency
maximum regression partial least squares, 398, 404–6
moving window PCA, 248–9
correlation matrix recalculation, 251–2
partial least squares, 395
computationally efficient algorithm, 394–6
confidence limits, 78–9, 119, 232
constraint NIPALS, 191–2, 192, 196
contribution charts, 95–8, 134–7
distillation example, 149–50
Q statistic, 97
reliability, 97–8
T2 statistic, 95–6
control ellipse, 17, 20–1, 30, 84–5, 304
construction, 23–4
control limits, 6, 8–9
definition, 11
monitoring statistics, 256–7

© 2012 John Wiley & Sons, Ltd. Published 2012 by John Wiley & Sons, Ltd.
control limits (continued)
  moving window principal component analysis (MWPCA), 252–3
scatter chart, 17
small data sets, 232–3
specification limits, 10
correlation see variable correlation
correlation coefficient, 15, 52–4, 146
  highly correlated variables, 19–25
  hypothesis testing, 25–6
  perfect correlation, 17–19
  uncorrelated variables, 16–17
correlation matrix, 17, 69–70, 243–4, 249, 281, 286, 334–5
chemical reaction process example, 126–7
moving window PCA, 244–7, 245–7
computational adaptation cost, 251–2
covariance
  highly correlated variables, 19–25
  perfectly correlated variables, 17–19
  primary residuals, 313
covariance matrix, 243, 260
adaptive PLS, 286
distillation process example, 144–5
effect of variable reconstruction, 116–17
eigenvalues, 21–3
partial least squares, 84–5
  estimation, 44
  MRPLS and, 60–1
primary residuals, 344–5
principal component analysis, 36–7
  error, 171–2
t-score, 84–5
reconstruction from loading vectors, 373–4
t-scores, 84
covariance structure changes, 293–4
examples, 294–7, 299–304
detectable by standard methods, 296
not detectable by standard methods, 296–7
fault isolation, 324–8
gearbox system example, 331–41
monitoring statistics, 322–3
Q statistic, 302
simulated examples, 294–304, 317–24, 328–31
$T^2$ statistic, 297–8, 301–2
cross-correlation matrix, 285–7, 290–291
cross-covariance matrix, 43, 46–7, 55
adaptive PLS, 286
deflation procedure, 216
primary residuals, 345–6
cross-validation, 76–7, 236
bootstrapping, 77–8, 287
leave-one-out (LOO), 146–7, 235–7
principal component analysis, 70–5
R statistic, 71–2
cumulative distribution function, 8
cumulative percentage variance (CPV), 67, 68, 69, 75
cumulative sum charts, 12
data matrix
  contribution of score vectors to, 364–5
  primary residuals, 376–7
deflation, 386–7
data reduction techniques, 28–9
dual-block, 29
partial least squares, 38–9
single-block, 28–9
effect of error-in-variable data structure, 216
MRPLS, 400–401, 405–6
distillation process example, 39, 40, 49–50, 141–3
error-in-variable model for maximum redundancy PLS, 223–6
fault identification, 153–63
column temperature, 157
feed flow drops, 153–7
monitoring model identification, 144–53
covariance matrix, 144–5
input variables, 146
MRPLS model identification, 146–7
divide and conquer algorithm, 250
drift faults, 328, 349
eigendecomposition, 66, 168–71, 199, 244, 363
changes in covariance structure, 318, 321, 335
determination of primary residuals, 306–7
effect on monitoring statistics, 118–19
error-in-variable estimation, 221
maximum likelihood PCA, 177–8
moving window PCA, 247–51
NIPALS algorithm, 192
recursive PCA, 281, 305
simulation example, 171
embedding vectors, 70
error covariance matrix, 258–9
error-in-variable data structure, 167, 214–17
distillation process example, 217–26
known error covariance matrices, 218–22
examples
fluid catalytic cracking see fluid catalytic cracking
gearbox system see gearbox system example
simulation see simulation examples
F-distribution, 85, 93, 130–131, 199–201, 235, 251
fault detection, 83–4, 94
gearbox, 338–41
improved residuals, 338–41
ramp changes, 258–61
residuals-based tests, 98–100
simulation example, 328–30
scatter diagrams, 84–5
fault diagnosis, 95
fault diagnosis (continued)
  chemical plant example, 133–8
  distillation process example, 153–63, 158–9
  using primary residuals, 324–8
fault direction, 325, 327
fault identification, 93–5
  improved residuals, 328
  variable reconstruction, 100–10
fault isolation, 94
  primary residuals, 324–8
feed flow, 153–5
first order perturbation, 248
floating point operations, 248–9, 251–2
Fluid Catalytic Cracking (FCC), 265–6, 266
  input sequences, 267
  loss in combustion air blower, 270–2
  process description, 266–8
  process variables, 268
  simulated examples, 268–74
fluidization (catalyst), 124
Frobenius norm, 55, 63–4
furnace process, 278–86
Gaussian distribution, 6–8, 7
  gearbox system example, 333
  monitoring model, 334–5
  process description, 332
  schematic, 333
  tooth breakage, 332–3
gradient descent, 248
Gram-Schmidt orthogonalization, 249
H-principle, 76, 79
Hotelling $T^2$ statistic, 83, 85, 87–9, 259, 263–4
  changes in covariance structure
  undetectable by, 302
  chemical process example, 130–133
  contribution charts, 95–6, 134–7, 135–6
distillation process example, 158
  effect of covariance structure
  changes, 295, 297–8
  effect of variable reconstruction, 118–19
  F-distribution, 85, 93, 130–1, 199–201
  improved residuals, 352–3
  maximum likelihood PCA, 198–202
MLPCA, 200
  moving window PCA, 256–7, 285
  partial least squares model, 91–2
  primary residuals, 316–17, 321–3
  ramp error detection, 260–1
  small data sets, 233–4, 236–7
  hypothesis testing, 12–14, 24–5, 88–9, 90, 92, 96–7, 99–100, 188, 195
  correlated variables, 24–6
  distillation example, 158–9, 160–3
  small data sets, 234–5
  Type I error, 13–14, 237, 262–3, 297
  Type II error, 14
see also $T^2$ statistic, $Q$ statistic
ideal process, 11, 15
  improved residuals, 294, 322–7, 342
  covariance matrix
    change in eigenvectors, 346–7
    change in score variables, 346
    for eigenvalues, 315–17
  gearbox system example, 331–41
  sensitivity to sensor bias, 349–53
  indicator function, 70
  inverse iteration, 248
Isserlis theorem, 204, 313, 316, 345
Kernel algorithm, 379
Kronecker delta, 32

Lanczos tridiagonalization, 248
latent variables (LV), 28, 42–3, 55, 76, 202–3, 210, 360
adaptive PLS, 290–2
distillation process example, 146–8
estimation, 46–8, 63–5, 223
see also stopping rules
retention stopping rules, 65–75, 290
least median of squares (LMS), 230–1
least trimmed squares (LTS), 231
leave-one-out (LOO) cross-validation, 146–7, 235–7
Liapounoff theorem, 367–8
loading matrix, 358
loading plots, 132, 149
loading vectors, 34, 38, 55–7, 149–50, 167, 215, 362–3, 376
computation, 379
error-in-variable data structure, 214, 215–16
distillation process example, 223–4
producing score variables of maximum variance, 190–194
reconstruction of covariance matrix, 373–4
small data sets, 233
see also p-loading; q-loading
M-estimator, 230
matrix deflation see deflation procedure
maximum likelihood principal component analysis (MLPCA), 66, 172–6
eigenvectors, 198–9
examples, 176–87
Hotelling T^2 statistic, 200
model and residual subspace estimate properties, 187–94
PLS error-in-variable, 218
simulated examples, 182–7, 194–202
application to data, 193–4
stopping rules, 187–8
unknown error covariance matrix, 177–82
maximum redundancy partial least squares (MRPLS), 54–6, 58–65
algorithm, 56
computationally efficient algorithm, 398
distillation example, 161–2
error-in-variable model, 223–6
process variables, 143
error-in-variable data structure, 215–17
distillation process example, 161–2
trend, 9–10
means squared error, 148
measurement bias, 226–7
minimum covariance distance estimator (MCD), 232
Minimum Description Length, 65–6
minimum volume estimator (MVE), 231, 231–2
model estimation see deflation procedure
model subspace, 31, 35, 306–7
estimation, 177–82
estimate properties, 189–94
simulation example, 34–8, 182–7
monitoring statistics, 85–6, 158, 232–3
covariance structure change sensitivity, 301–2
maximum regression PLS (MRPLS), 92–3
moving-window PCA, 252–3
partial least squares (PLS), 90–92
primary residuals, 352–3
principal component analysis (PCA), 86–90
Monte Carlo simulations, 207–8, 210, 234, 301–2, 318, 321
moving window partial least squares (MWPLS), 242
moving window principal component analysis (MWPCA), 241–2, 328
application delay, 253–4, 276–8
correlation matrix, 244–7
downdating, 245–6
updating, 246, 247
eigendecomposition, 247–51
example, 282–6
model determination, 275–6
simulated example, 261–5
fluid catalytic cracking, 275–8
simulation example, 258–65
source signal determination, 250–1
window length, 254–7
multivariate trimming (MVT), 231
NIPALS, 56, 379
see also constraint NIPALS
non-negative quadratic monitoring statistics see monitoring statistics
null hypothesis, 13, 90
ordinary least squares (OLS), 46, 53, 203–4, 217, 393–4
compared to PLS, 205–12
parameter estimation, 207
regression parameters, 207–9, 208
outliers, 167–8, 226–7
trimming, 231–2
p-loading vectors
orthogonality to w-weight vectors, 389–90
orthonormality, 366
see also loading vectors
parallel analysis, 70
parameter estimation, 47–9
least median of squares (LMS), 230–231
M-estimator, 230
ordinary least squares, 205–12
partial least squares, 202–26, 212–15, 219–22
distillation process example, 223–6
projection pursuit, 230
robust estimation of moments, 228–9
small sample sets, 232–7
trimming, 231–2
Pareto Maxim, 4
partial least squares (PLS), 38–49
adaptive, 286–90, 286–92
model adaptation, 290–291
algorithm overview, 375–6, 380
compared to ordinary least squares (OLS), 205–12
computational efficiency, 394–5
contrasted with maximum redundancy PLS, 60–65
core algorithm, 377–9
data structure assumptions, 39–41, 49–50
deflation procedure, 41–3
distillation process example, 147–8
error-in-variable structure, 212–15
loading vectors, 42
maximum redundancy, 49–58
model identification, 51–2
model properties, 381–2
matrix-vector product properties, 384–6
regression coefficient calculation, 390
relationship between q-weight and q-loading vectors, 391
t- and u-score vector orthogonality, 385–6
t-score vector calculation from data matrix, 388
t-score vector orthogonality, 382–4
w-weight to p-loading vector orthogonality, 389–90
monitoring statistics, 91–5
non-negative quadratic process monitoring, 86–90
partial least squares, 90–93
parameter estimation, 202–26
parameter estimation bias and variance, 203–5
regression model accuracy, 393–4
score variables, 41
simulation example, 43–9
input variable set, 43–4
latent variables, 45–6
PLS model determination, 44–5
weight and loading matrices, 46–7
simulation examples, limitations, 58–65
stopping rules, 76–9
bootstrapping, 77–9, 287
variable sets, 38–9
weight vectors, 41
see also maximum likelihood partial least squares
Powerforming, 278
prediction sum of squares (PRESS), 71–2, 146–7, 288–90, 290–291
predictor variable sets, 38–9
primary residuals, 294
covariance, 313
covariance matrix, 344–5
change in score values, 347–8
eigenvectors non-orthogonal to model, 348–9
degrees of freedom, 308–9
distribution function, 313–14, 320, 323
eigenvectors, 306
fault isolation, 324–8
gearbox example, 341–53
residual subspace, 294, 305–7
sensitivity, 310
simulation examples, 318–24
statistical properties, 312–15
variance, 313, 319
principal component analysis (PCA), 30–38
algorithm summary, 362–3, 380
chemical process example, 125–33
computation
principal component analysis (PCA) (continued)
core algorithm, 357–62
scaling matrix, 357–8
covariance matrix, 36–7
data correlation matrix, 334–5
data structure assumptions, 30–33, 49–50, 86
eigendecomposition of covariance matrix, 168–70
Fluid Catalytic Cracking example, 273–5
geometric analysis of data structure, 33–4
limitations regarding time-varying processes, 280–282
loading vectors, 34
model identification, 50–51, 126
model properties, 363–74
asymptotic distribution of t-score, 366–9
covariance matrix exhaustion, 372–3
data matrix exhaustion, 372
orthogonality of t-score vectors, 365
orthonormality of p-loading vectors, 366–7
t-score vector computation, 371–2
non-negative quadratic process monitoring, 86–90
residual subspace, primary residuals, 306–8
residuals-based tests, 98–100
robust parameter estimation, 230–1
simulated examples, 258–61
simulation example, 34–8
stopping rules
cross-validation based, 70–75
eigenvalue-based, 67–9
information-based, 65–7
time-varying processes, 241–2, 258–61
see also moving window principal component analysis, maximum likelihood principal component analysis
probability density function, 6–8
improved residuals, 315
perfectly correlated variables, 18–19
uncorrelated variables, 16–17
process monitoring, 81–2
see also monitoring statistics
process types, 10–11
projection pursuit, 230
projection-based adaptation, 248
promising process, 11
propane, 141
Q statistic, 83–4, 89–90, 104–5, 130–133, 258–9
changes in covariance structure undetectable by, 297, 302–4
contribution charts, 97, 134–7, 135–6
effect of variable reconstruction, 118–19
maximum likelihood PCA, 198–202
moving window PCA, 252–3, 256–7, 285
primary residuals, 316–17, 322, 353
small data sets, 234–5, 236–7
q-loading, 150, 391
see also loading vectors
q-weight, 46–7, 57, 391
maximum redundancy PLS, 404, 406–9
quadratic monitoring statistics see monitoring statistics
R statistic, 71–2
r-weight vectors, 387–8
ramp error detection, 258–65
random variables
  cumulative distribution function, 8
  mean and variance, 5–6, 9–10
  probability density function, 6–8
  Shewhart charts, 8–9, 12
  trends, 9–10
rank one modification, 248
recursive principal component analysis (RPCA), 241
regression coefficient, 390
residual percentage variance test, 69
residual subspace, 31
  effect of variable reconstruction, 118
  estimation, 186–7
  estimate properties, 189–94
  simulation examples, 34–8
fault detection based on, 98–100
maximum likelihood PCA, 186–7, 187–94
primary residuals, 294, 305–7
  sensitivity, 310
  statistical properties, 312–15
see also improved residuals;
  primary residuals
residual sum of squares (RSS), 71–2
RSS, 70–1

sample generation, 329
scaling matrix, 357–8
scatter diagram, 16–17, 82
  weaknesses, 296–7
scatter diagrams, 84–5
  distillation example, 158, 159–60, 161
score variables, 41–2, 362
  weaknesses, 296–8
see also t-score, u-score
score vectors, contribution to data matrix, 364–5
SCREE test, 69, 70
sensor bias, 90, 102, 108, 279, 281, 285
  sensitivity of improved residuals, 349–53
Shewhart charts, 8–9, 10, 12, 15–20, 22–3, 232
significance, 8
SIMPLS algorithm, 379
simulation examples
  adaptive MSPC, 257–61
  covariance structure changes, 294–304, 317–24, 328–31
Fluid Catalytic Cracking, 268–74
  PCA application, 273–5
maxumum likelihood PCA, 182–7
model and residual subspace estimation, 34–8
partial least squares (PLS), 43–9, 58–65
weaknesses of conventional MSPC, 319–24
singular value decomposition (SVD), 174, 359–60, 382
source signal adaptation, 287–8
source signal determination, 194–5
time variant, 250
special cause variation, 11–12, 31
specification limits, 10–11, 10
squared prediction error see Q statistic
Stahel-Donoho location estimator, 229
standard deviation, 6
statistical fingerprinting, 81–2
statistical local approach, 304–5, 318–20, 341–2
statistical process control (SPC)
  basic principles, 5–12
  history, 4–5
statistical process control (SPC)  
(continued)  
motivations for use of  
multivariate techniques,  
15–24  
overview, 3–4  
step faults, 90, 102, 325–8  
stochastic variables see random variables  
stopping rules  
maximum likelihood PCA  
(MLPCA), 187–8  
partial least squares (PLS), 55  
adaptive, 287  
analysis of variance, 76  
bootstrapping, 77–9, 287  
cross-validation, 76–7  
principal component analysis  
(PCA)  
cross-validation based,  
70–75  
eigenvalue-based, 67–9  
information-based, 65–7  
Student t-score see t-score  
t-score, 42, 51–2, 83  
asymptotic distribution of  
variables, 367  
distillation process example,  
148–50, 152–3, 152, 160, 161, 162  
error-in-variable data structure,  
224  
maximum regression PLS,  
397–8  
MRPLS models, 92–3  
partial least squares models, 210  
asymptotic distribution,  
391–2  
computation, 380  
orthogonality, 381  
orthogonality with u-score,  
385–6  
vector computation, 371–2  
vector orthogonality, 365–6,  
382–3, 397  
see also score variables  
t’-score, 83, 93  
distillation process example, 152, 160  
see also score variables  
time-varying processes, 241–2  
application delays, 253–4,  
261–5, 275–8, 282–6  
minimum window length,  
256–7  
partial least squares methods,  
287–92  
source signal determination,  
250–251  
tooth breakage, 332–3  
total quality management (TQM),  
5  
treacherous process, 11  
trends in mean value, 9–10  
see also drift faults  
trimming, 231–2  
turbulent process, 11  
Type I errors, 234, 263–4  
Type II errors, 25–6, 25  
u-score, 53–4, 376  
maximum regression PLS,  
406–9  
orthogonality with t-score  
vectors, 385–6, 397–8  
see also score variables  
variable correlation, 15–16  
perfect correlation, 17–19  
uncorrelated variables, 16–17  
see also correlation coefficient;  
correlation matrix  
variable reconstruction  
chemical process example,  
126–7  
influence on model plane,  
117–18  
influence on residual subspace,  
118
projection-based, 100–3
geometric analysis, 112–13
limitations, 109
linear dependency of
projection residuals, 111–12
maximum dimension of
fault subspace, 115–16
optimality, 113–14
reconstruction subspace, 114–15
regression formulation, 104–5
single sample, 107–9
regression-based, 109–11
variable reconstruction error, 126, 128, 281–2, 335
variance, 6
primary residuals, 313
trends, 9–10
variance of reconstruction error
(VRE), 73–5, 262, 273–4
Velicer Partial Correlation
Correction (VPC), 75, 250, 252, 335
vibration, 334
W statistic, 72
w-weight, 46–7, 58, 62, 291, 380, 397–8, 401, 405–8
weight vectors, 41–2, 52, 55–6, 215, 376, 381–2
Wishart distribution, 233