**Reflections on ME 233, Spring 2012**

**General Notes**

* I divided the course into 4 units:

1. Probability review
2. State-space control and estimation
3. Input/output control
4. Adaptive control

By doing this, I hoped to give the students a better sense of the course structure and it allowed me to highlight throughout the course where we were in the course material.

* Throughout the course, I only discussed discrete-time material for the sake of consistency. Although most (if not all) of the techniques work in continuous-time, I felt that it would take too much time to do both.

**Unit 0: Probability review**

* To save time, I completely omitted the lecture on scalar random sequences.
* Some students found the notation versus confusing in 2011. I changed the notation so that the Z-transform of is .

**Unit 1: State-space control and estimation**

* I restructured this material so that all of the finite-horizon material would be covered before moving on to infinite-horizon material.
* Whenever it was easy, I allowed the LQR/LQG cost function to have a cross-term between the state and the control.
* It might be worth considering a notation change for the optimal control policy for the LQR and LQG control. In particular, when discussing LQG, the superscript ‘o’ is used to denote both the Kalman filter a priori estimates and the optimal control. Although it is reasonably easy to distinguish between the two types of signals (because estimates have hats), it might be more clear to change the notation for the optimal control signal to something else, e.g. .
* I changed the LQR and LQG cost functions so that they do not have a factor of 1/2. My thought is that the factor of 1/2 doesn’t serve any useful purpose and is just one extra thing to have to write. Moreover, the factor of 1/2 can always be absorbed into the cost function matrices.
* Due to a lack of time to write a problem, I didn’t assign any homework on FSLQR design. The only problem I asked that was related to FSLQR was regarding the existence conditions.

**Unit 2: Input/output control**

* Since Unit 2 is a collection of control design techniques, I made sure to explicitly say before starting each lecture where we were in that unit. In particular, I always mentioned which of the 4 techniques we had already finished, which one we were covering that day, and which one(s) would be covered in later lectures. My hope was that it would help the students keep track of what they’ve done and what was still left to do in this unit.
* Throughout this section, the term “monic” was misused to mean polynomials with constant coefficient equal to 1. I have fixed this.
* I used a different convention for referring to when a polynomial corresponds to a stable system. In particular, I said that if a polynomial has its zeros strictly inside the unit circle, then it is Schur; if it has its zeros strictly outside the unit circle, then it is anti-Schur.

Another possibility that I came up with was to use the terminology Schur to refer to functions instead of polynomials. That way, would be a Schur function of *z*.

* Throughout this unit, I replaced the superscript \* with a bar; I didn’t want people to get confused between taking the complex conjugate of a polynomial and multiplying by *zr* for some value of *r*.

**Unit 3: Adaptive control**

* The formula for recursive least squares with forgetting factor was incorrect. (The PAA presented in previous years achieves output error convergence, but shouldn’t be interpreted as a forgetting factor.) I changed all of the PAAs, proofs, etc so that is the forgetting factor.
* The conditions on λ1(*k*) and λ2(*k*) should be changed to:

If λ1(*k*) is not bounded away from 0, then *eo*(*k*) might not converge to 0. If λ2(*k*) is not bounded away from 2, then the LTI block 1- λ/2 cannot be made SPR, which means that the asymptotic hyperstability theorem cannot be used. Although I made this change in L25, it is inconsistent in the other presentations (L20-L24).

**Lectures**

|  |  |
| --- | --- |
| L1 | I was questioning whether or not to cover discrete random variables in this course. After all, we don’t use discrete random variables at all in the main material of the course. Eventually, I decided to cover them, but some time could be saved (if necessary) by skipping this. |
| L2 |  |
| L3 | Although I did the proof of the conditional PDF for Gaussians slightly differently than in the slides, I did not update the slides to reflect this. |
| L4 | The term “white” was abused in this lecture, i.e. it was used for non-WSS random vector sequences. I have replaced “white” by “uncorrelated” when referring to non-WSS random vector sequences in this lecture and the following lectures. (This appears to be the correct terminology.) Also, I believe the term “ergodic” is being used incorrectly in this lecture. The book “Random Processes: A Mathematical Approach for Engineers” by Robert M. Gray and Lee D. Davisson appears to have a good discussion of common misconceptions/misuses of that term. |
| L5 | I switched much of the notation to capital variables, i.e. I used estimators instead of estimates. My thought was that it is simpler to deal with estimators, which don’t require any modification of the PDF to do the analysis. Also, in this lecture, I removed some material (e.g. Law of Variances), so you will have to get that material from the Spring 2010 PowerPoint lectures if you want to cover that material. Regarding the proof of least squares property 3, I did the proof in a different manner than in the slides. However, I did not update the slides to use my approach. One consequence is that the slides for this proof still use the “estimate” approach rather than the “estimator” approach. |
| L6 | I did the proof of the Kalman filter on the board. There were two differences between how I did it and how it was presented in the slides. First, instead of examining k=1, I looked at general k>0 under the assumption that I had access to the relevant quantities from the previous step. Second, I explicitly listed which vectors depended on which random vectors. See Table 1 below. |
| L7 | I used the technique of completing the square to solve the optimization problems that arise in LQR. I later realized that this makes homework problems somewhat more difficult, so you might want to revert this proof to one based on derivatives. This year, I also presented on the board the proof based on derivatives. |
| L8 | I changed the proof of finite-horizon LQG so that there was no notion of taking the expectation over only a subset of the randomness in the system. Also, I found what seemed to be an error in the optimal output feedback LQG cost. You might want to double check what I came up with. |
| L9 | * I personally think that all of this material should be covered in ME 232, instead of just some of it. * In addition to the material in the presentation, I also presented a proof on the board of the equivalence of conditions 2 and 4 for observability. Essentially, I proved that:  1. If the spectral condition (condition 4) is not met, then *A* has an eigenvector in the nullspace of *C*. This eigenvector is in the null space of the observability matrix, which implies that the observability matrix does not have rank *n.* 2. I stated that if *x* is in the unobservable subspace, then *Ax* is also in the unobservable subspace. Thus, the unobservable subspace is *A*-invariant, which implies that it contains an eigenvector *y* of *A* if the unobservable subspace is nontrivial. Since the unobservable subspace is the null space of the observability matrix, it must hold that *Cy* = 0. Therefore, if condition 2 is not met, then there exists an eigenvector of *A* that is in the null space of *C*. This vector is in the null space of the matrix in the spectral condition, which implies that the spectral condition is not met.  * The final theorem was proven in homework 5, problem 1. |
| L10 | * Some students found the notes on Theorem 1 (existence of a bounded limiting solution) to be confusing. A rewording of these notes might be in order. * I think it should be highlighted that DAREs can have multiple solutions and that we are only interested in one particular solution of the DARE. In particular, depending on which set of existence conditions we are using, either the positive definite solution of the DARE, the positive semi-definite solution of the DARE, or the one that yields stable closed-loop dynamics. * Reading through the lecture again, I see that the optimal control law is not stated prominently. A slide needs to be added for this. * Homework 5, problem 2 investigated the special case of LQR with zero *S* and proved that the transmission zeros condition reduces to a condition involving unobservable modes. |
| L11 | I deemphasized the guaranteed LQR gain and phase margins. My thought is that, in order to compute the discrete-time guaranteed LQR gain and phase margins, you need to essentially perform the complete control design. From a practical standpoint, if you perform the complete control design, you probably are using MATLAB to solve the DARE, which means that you might as well also use MATLAB to compute the actual gain and phase margins. Moreover, there are more suitable techniques for robust control. Thus, there are very few reasonable uses for the guaranteed phase and gain margins of discrete-time LQR in this day and age. |
| L12 | * Although the slides are hidden for the interpretation of the return difference equality (when *A* is Schur), I did cover it on the blackboard. (When A is not Schur, it doesn’t make sense to talk about the power spectral density of *Y*.) * This lecture should include more details for the reciprocal root locus. |
| L13 | In addition to the material in the slides, I talked about HDD control. I mentioned LQG with variance constraints, but didn’t give any formal treatment of it; I just mentioned how it related to choosing the weights in the control design process. (See HW6, problem 1) |
| L14 | The set of sufficient conditions for FSLQR given on slide 31 was proven in hw 6, problem 2. |
| L15 |  |
| L16 | I removed the factoring convention for *B(q-1)* from the lecture. This way the students can choose *Bs* and *Bu* in whatever way is most convenient for the problem. Also, I explicitly proved that an *n*th order polynomial is Schur if and only if the polynomial is anti-Schur (even if has zeros at the origin). This is not in the slides. |
| L17 | Again, I removed the factoring convention for *B(q-1)* from the lecture. Also, I presented the internal model principle on the board. (It is not in the slides.) |
| L18 | * To prevent confusion, I used hats for all matrices in this lecture. * It should be mentioned that w(k) and v(k) (for the state space system) and ε(k) (for the ARMAX model) are Gaussian * It might be helpful to have block diagrams to show the definitions of the various signals |
| L19 | * According to the Astrom and Wittenmark [TODO: fix] book, STR and MRAS are the same. Is it necessary to mention them in the introduction to this lecture? * I removed the Kalman-Szego-Popov lemma and instead replaced it by a strict matrix inequality condition. (In particular, it seemed that the condition you had was incorrect. At the very least, the realization of the system should be minimal to use the result that you stated.) When I presented this result, I mentioned on the blackboard the similarity between that condition and the Lyapunov stability result (i.e. there exists such that the condition holds). * One subtle detail in the slides is that when the sufficiency part of the hyperstability theorem is proven, the squared norm of *xk* is bounded, rather than its norm. |
| L20 | Due to a lack of time, I moved some of the details to the end of the lecture. However, time permitting, it would be a good idea to cover the topics in the “Additional Material” section during lecture. |
| L21 | * I included the definition in this lecture. This makes the role of much more clear. * The conditions on λ1(*k*) and λ2(*k*) should be made consistent with L20. |
| L22 | I significantly revamped this presentation. In the process, I ended up cutting a lot of material out (i.e. I hid the slides). One thing I meant to change in the slides, (but forgot) is that instead of saying that colored noise is PE of any order, I think the result should be that any sample sequence of colored noise is PE of any order. If you use the non-sample-sequence version, everything is random, including *Cn*, which doesn’t make sense, at least to me. (Although I admit that I could be wrong about this.) |
| L23 | * I flipped the factoring convention for *B*; when all of the zeros are stable, we choose *Bu* = 1. This way, the case for which *B*(*q*-1) is anti-Schur is a special case of the general algorithm. * It might be worthwhile to add a slide that discusses how to factor *B* in an automated fashion. * The conditions on λ1(*k*) and λ2(*k*) should be made consistent with L20. |
| L24 |  |
| L25 | * I think that color-coding the proof that the relevant nonlinearity is P-class was helpful. * I tried a new idea out in this lecture: I put a horizontal line below any material that was just being copied from the previous slide. (See Part 3 of the stability proof for examples.) |

Table 1. Stochastic Dependence for L6

|  |  |
| --- | --- |
| **Random vector** | **Dependence on x(0), w, and v** |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**Homework**

|  |  |
| --- | --- |
| 1 |  |
| 2 |  |
| 3 |  |
| 4 |  |
| 5 | * Problem 1b could be worded better to highlight that we want to show that the first bullet is equivalent to the second bullet. * I significantly shorted problems 3 and 4. |
| 6 | There are a number of ways in which problem 1 can be improved.   * Some students misread 3\sqrt as the cube root. Maybe this should be stated in terms of 3σ. * Part c of the hint says that “current” should be an input into the lqgreg function. However, it can instead be used as an input to the kalman function or as an input to both function. Any of these three approaches will work. However, MATLAB will give a warning (not an error) when “current” is used as an input argument to the lqgreg function. |
| 7 | Problem 1b is somewhat misstated. In particular, it should ask if there exists real α such that the control law is implementable and the closed-loop transfer function from *D(z)* to *P(z)* is zero. Also, it would probably be appropriate to mention some condition on Δ(*z*) so that you can take the limit as α approaches 1 when analyzing the transfer function. Finally, the question should lead the students to realize that choosing α close to 1 causes the open-loop transfer function from to *U*(*z*) to become large. In other words, there is a tradeoff between disturbance rejection and the size of the control signal. |
| 8 |  |
| 9 | * I modified problems 3 and 4 so that the PAA agrees with the one presented in lecture. * I modified sp\_predict.m so that it preallocates memory for the vectors and so that the PAA agrees with the one presented in lecture. * I modified the proof of problem 4 so that it does not require *C(q-1)* to be anti-Schur. |

**Exams**

(See the pdf files for comments)